

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2006, 01:28:24 ; Search time 189 Seconds
(without alignments)
1011.269 Million cell updates/sec

Title: US-10-717-619-2
Perfect score: 2316
Sequence: 1 MTSQGNKRTTKEGFDLRFQ.....QPLPLARLLLTQYSSQALHE 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	435	4 AAB81070	Aab81070 Murine ma
2	1316.5	56.8	428	5 AAE22609	Aae22609 Human MIS
3	1316.5	56.8	428	7 ADB65190	Adb65190 Human pro
4	1316.5	56.8	443	5 AAE22608	Aae22608 Human MIS
5	1132	48.9	376	4 AAB81071	Aab81071 Human mas
6	1109.5	47.9	353	5 AAE22610	Aae22610 Human MIS
7	398	17.2	533	8 ADP65191	Adp65191 Human lym
8	398	17.2	533	8 ADP23847	Adp23847 Human sof
9	398	17.2	533	8 ADP23845	Adp23845 PRO poly
10	398	17.2	533	8 ADP23845	Adp23845 PRO poly
11	398	17.2	533	8 ADP23845	Adp23845 PRO poly
12	337	14.6	224	8 ADP45459	Adp45459 Human col
13	337	14.6	225	8 ADP45461	Adp45461 Human col
14	337	14.6	230	8 ADP45460	Adp45460 Human col
15	326	14.1	456	4 AAB93406	Aab93406 Human pro
16	319	13.8	449	2 AAW59866	Aaw59866 Amino aci
17	319	13.8	456	2 AAW59865	Aaw59865 Amino aci
18	319	13.8	456	9 ADZ26491	Adz26491 Human BLI
19	287.5	12.4	457	2 AAW59867	Aaw59867 Amino aci
20	260	11.2	95	2 AAW59867	Aaw59867 Human SLP
21	213	9.2	84	3 AAG03152	Aag03152 Human sec
22	212	9.2	46	4 AAM18973	Aam18973 Peptide #
23	212	9.2	46	4 ABB38147	Abb38147 Peptide #
24	212	9.2	46	4 AAM31575	Aam31575 Peptide #

25	212	9.2	46	4 ABB23354	Abb23354 Protein #
26	212	9.2	46	4 AAM71296	Aam71296 Human bon
27	212	9.2	46	4 AAM58783	Aam58783 Human bra
28	212	9.2	46	4 ABG53007	Abg53007 Human liv
29	212	9.2	46	5 ABG41096	Abg41096 Human pep
30	185.5	8.0	503	4 ABG22854	Abg22854 Novel hum
31	185.5	8.0	503	7 ADC32915	Adc32915 Human nov
32	185.5	7.1	474	2 AAM15253	Aam15253 Human bra
33	185.5	7.1	594	2 AAM15256	Aam15256 Human bra
34	165.5	7.1	594	9 ADX05533	Adx05533 Cyclin-de
35	165.5	7.1	594	9 ADZ28263	Adz28263 Human SHC
36	161.5	7.0	344	7 ADE14373	Ade14373 Human int
37	150.5	6.5	743	4 AAM79738	Aam79738 Human pro
38	146	6.3	728	7 ADD46841	Add46841 Human pro
39	146	6.3	728	7 ADE56294	Ade56294 Human pro
40	146	6.3	728	8 ADJ66562	Adj66562 P13 kinas
41	144.5	6.2	474	2 AAM39087	Aam39087 Human SHC
42	144.5	6.2	474	2 AAM39089	Aam39089 Human SHC
43	143.5	6.2	722	7 ADD46839	Add46839 Rat Prote
44	143.5	6.2	722	7 ADE56292	Ade56292 Rat Prote
45	143	6.2	864	8 ABO84879	Abo84879 Murine ca

ALIGNMENTS

RESULT 1
AAB81070
ID AAB81070 standard; protein; 435 AA.
XX
AC AAB81070;
XX
DT 25-JUN-2001 (first entry)
XX
DE Murine mast cell-specific signal transduction protein.
XX
KW Mast cell; signal transduction; mouse; allergic disease.
XX
OS Mus musculus.
XX
PN JP31146204-B13
PD 12-MAR-2001.
XX
PF 17-SEP-1999; 99JP-00263778.
XX
PR 17-SEP-1999; 99JP-00263778.
XX
PA (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.
XX
DR WPI; 2001-310022/33.
DR N-PSDB; AAF86139.
XX
PT Mast cell-specific signal-transduction molecule, useful for screening
therapeutic compounds for treating allergies, is specifically expressed
by mouse mast cell.
XX
PS Claim 1; Page 7-8; 12pp; Japanese.
XX
CC This sequence represents a murine mast cell-specific signal transduction
protein. The invention includes the cDNA and protein sequences of the
mast cell-specific signal transduction molecule and an expression vector
containing the polynucleotide sequence. The coding sequence of the signal
transduction protein can be used for screening therapeutic compounds
which will be useful for treating allergic diseases
XX
SQ Sequence 435 AA;

Query Match 100.0%; Score 2316; DB 4; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.9e-193;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTSQGNKRTTKEGFDLRFQNVLSLLKNRSWPSLSAKGRCAVLEPLPDHRRNLAVPGG 60

Db 1 MTSGGNKRTTKEGFGDLRFQNVSLKNRSPWPSLSAKGRCAVLPLEPDRHRLNLAGVPGG 60
QY 61 EKNSNNNDYEDPEFQLLKAWPSMKILPARPIQSEYADTRYFQDMMAPLILPPKASVST 120
Db 61 EKNSNNNDYEDPEFQLLKAWPSMKILPARPIQSEYADTRYFQDMMAPLILPPKASVST 120
QY 121 EROTRDVRMTQLEBEVDKPTFKDVRQSRQFSGFKYTKINKTKPLPPRPAITILPKKYQPLPPA 180
Db 121 EROTRDVRMTQLEBEVDKPTFKDVRQSRQFSGFKYTKINKTKPLPPRPAITILPKKYQPLPPA 180
QY 181 PPESSAYFAPKPTFFPEVQGRQPRQSAKDFSRVLGAEEESHQTKPESSCPSSNQNTOKS 240
Db 181 PPESSAYFAPKPTFFPEVQGRQPRQSAKDFSRVLGAEEESHQTKPESSCPSSNQNTOKS 240
QY 241 PPAIASSSYMPGKHSIQARDHTGSMQHCPCQAORCOAAASHSPRLPYENTNSEXPDPTKPD 300
Db 241 PPAIASSSYMPGKHSIQARDHTGSMQHCPCQAORCOAAASHSPRLPYENTNSEXPDPTKPD 300
QY 301 EKDVQWQNEWYIGEYSRQAVEDVLKMKENKDGTFVLVDCSTKSKAEPYVLVVFYGNKVNVK 360
Db 301 EKDVQWQNEWYIGEYSRQAVEDVLKMKENKDGTFVLVDCSTKSKAEPYVLVVFYGNKVNVK 360
QY 361 IRFLESNQOQFALGTGLRGNMFDSDVEDIIEHYTYFPILLIDGDKKAARRKQCYLTQPLPL 420
Db 361 IRFLESNQOQFALGTGLRGNMFDSDVEDIIEHYTYFPILLIDGDKKAARRKQCYLTQPLPL 420
QY 421 ARLLLTQYSSQALHE 435
Db 421 ARLLLTQYSSQALHE 435

RESULT 2
AAE22609
ID AAE22609 standard; protein; 428 AA.
XX AC AAE22609;
XX AC AAE22609;
DT 26-JUL-2002 (first entry)
DE Human MIST splice variant protein from clone #7.
KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
XX OS Homo sapiens.
XX PN WO200226986-A2.
XX PD 04-APR-2002.
XX 28-SEP-2001; 2001WO-US030593.
XX PF 29-SEP-2000; 2000US-0237030P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX WPI; 2002-372126/40.
DR N-PSDB; AAD35801.
XX New isolated mast cell immunoreceptor signal transducer polypeptide,
PT useful for treating immune disorder involving hyperactivity of B- or T-
PT lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma.
XX Claim 11; Fig 5; 171pp; English.
XX The present invention relates to novel mast cell immunoreceptor signal
CC transducer (MIST) proteins and polynucleotides encoding such proteins.
CC MIST sequences of the invention are useful for preventing, treating or
CC ameliorating a medical condition in mammalian subject. They are useful

CC for treating an immune disorder involving hyperactivity of B- or T-
CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful as
CC targets for therapeutic intervention in immune cell disorders and
CC inflammatory indicated, for diagnosis and/or screening of disorders or
CC diseases associated with expression of MIST, for screening for
CC antagonists or inhibitors of the interaction of MIST with cellular
CC signalling components. They are used in assays that detect activation or
CC induction of various B and T-cell-related neoplasms or cancers. Sequences
CC of the invention are also used in gene therapy. The present sequence is
CC human MIST splice variant protein from clone #7
XX SQ Sequence 428 AA;
Query Match 56.8%; Score 1316.5; DB 5; Length 428;
Best Local Similarity 62.1%; Pred. No. 4.4e-106;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;
QY 1 MTSGGNKRTTKEGFGDLRFQNVSLKNRSPWPSLSAKGRCAVLPLEPDRHRLNLAGVPGG 60
Db 1 MNRQGNKRTTKEGFGDLRFQNVSLKNRSPWPSLSAKGRCAVLPLEPDRHRLNLAGVPGG 60
QY 61 EKNSNNNDYEDPEFQLLKAWPSMKILPARPIQSEYADTRYFQDMMAPLILPPKASVST 120
Db 61 AKGHSDDDDYDDPELMEETWQSIKILPARPIQSEYADTRYFQDMMAPLILPPKASVST 120
QY 121 EROTRDVRMTQLEBEVDKPTFKDVRQSRQFSGFKYTKINKTKPLPPRPAITILPKKYQPLPPA 180
Db 121 GQPTWNTQ-TRLERVDKPTSKDVRQSRQFSGFKYTKINKTKPLPPRPAITILPKKYQPLPP- 178
QY 181 PPESSAYFAPKPTFFPEVQGRQPRQSAKDFSRVLGAEEESHQTKPESSCPSSNQNTOKS 240
Db 179 EPESRRPPLSQRHTFFPEVQRMPSQISLRDLSEVLEAEKVPNHQKPESTHLENQNTQEI 238
QY 241 PPAIASSSYMPGKHSIQARDHTGSMQHCPCQAORCOAAASHSP--RMLPYENTNSEXPDPTK 298
Db 239 PLAISSTTSNHSVQNRDHRGMPQPCSPQRCPPASCSPHENILPKYKTSWRPPPKR 298
QY 299 PDEKDVQWQNEWYIGEYSRQAVEDVLKMKENKDGTFVLVDCSTKSKAEPYVLVVFYGNKVNV 358
Db 299 SDRKDVQWQNEWYIGEYSRQAVEDVLKMKENKDGTFVLVDCSTKSKAEPYVLVVFYGNKVNV 358
QY 359 VKIRFLESNQOQFALGTGLRGNMFDSDVEDIIEHYTYFPILLIDGDKKAARRKQCYLTQ 417
Db 359 VKIRFLESNQOQFALGTGLRGNMFDSDVEDIIEHYTYFPILLIDGDKKAARRKQCYLTQ 418
QY 418 LPLARLLL 425
Db 419 LPLTRHLL 426
RESULT 3
ADB65190
ID ADB65190 standard; protein; 428 AA.
XX AC ADB65190;
XX AC ADB65190;
DT 04-DEC-2003 (first entry)
DE Human protein encoded by clone SPLEN20191020.
KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW cell regeneration; membrane protein; signal transduction-related protein;
KW transcription-related protein; osteoporosis; neurological disease;
KW cancer; tumour.
XX OS Homo sapiens.
XX PN EP1308459-A2.
XX PD 07-MAY-2003.
XX 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.
PR 25-JAN-2002; 2002US-00350978.
XX (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI: 2003-450961/43.
DR N-PSDB; ADB63220.
XX New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX Claim 1; Page; 222pp; English.
XX The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesizing the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours). The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the invention. Note: Some of the
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX SQ Sequence 428 AA;
Query Match 56.8%; Score 1316.5; DB 7; Length 428;
Best Local Similarity 62.1%; Pred. No. 4.4e-106;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;
QY 1 MTSQGNKRTTKGFGDLRFQNVLLKNRSPSSAKGRCAVLEPLDHRRLAGVPGG 60
Db 1 MNRQGNKRTTKGSDNLFQNFSLPKRNSPRINSATGQYQRMNKPLLDWERNFAAVLDG 60
QY 61 EKCSNNNDYDEPQLKAWPMKILPAPIQEYADTRYFQDMMEAPILLPKASVST 120
Db 61 AKGHSDDYDPELRMEETWQSIKILPAPIKESYADTRYFKVAMDTPLDTRTSISI 120
QY 121 BRQTRDVRMTQLEVDYPTKDVRSQRFKGYTKINKTLPPLPPRAITLPKKYQPLPPA 180
Db 121 GOPTWNTQ-TGLEKVDKPIKDVRSQIKGDASVRKNKILPPLPPRPLITLPKKYQPLPP- 178
QY 181 PPESSAYFAPKTPFPVQRPQRORSAKDFSRVLGAEEESHQPKPSSCPSSNQTKS 240
Db 179 EPSSRRPLSQRHPTFPVQRPQSIQLRDLSEVLEAKVPHNQKRPSTHLENQNTQEI 238
QY 241 PPAIASSYMPGKHSIQARDHTGSMQHCQAQRCQAAASHSP--RMLPYENTNSKPDPTK 298
Db 239 PLAISSSFTTSNHSVQNRDRHGGQPCSPQRCQPPASCSPHENILPKYKTSWRPPFKR 298
QY 299 PDEKDVQWQNYIGEYSRQAVEDVLMKENKDGTFPLVDCSTKSAEYPVLVFGNKVYN 358
Db 299 SDRKDVQWQNYIGEYSRQAVBEAFMKNKDGSLVDCSTKSKKEYPVLVAVFENKVYN 358
QY 359 VKTRFLESNOQFALGTGLRGNEMPDSVEDIEHYTYPILLIDGKDK-AARRKQCYLTOP 417
Db 359 VKTRFLESNOQFALGTGLRGNEMPDSVEDIEHYTYPILLIDGKDKTGVRHKKCHLTOP 418
QY 418 LPLARLL 425
Db 419 LPLTRHLL 426
RESULT 4
AAE22608
ID AAE22608 standard; protein; 443 AA.
XX AAE22608;
XX 26-JUL-2002 (first entry)
XX Human MIST protein #1.
XX Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Modified-site 84
FT /label= Tyrosine phosphorylation_binding_site
FT Modified-site 111
FT /label= Tyrosine phosphorylation_binding_site
FT Domain 306..311
FT /label= SH3_binding_proline-rich_motif
FT Domain 324..407
FT /label= SH2_domain
XX WO200226986-A2.
PN 04-APR-2002.
XX 28-SEP-2001; 2001WO-US030593.
XX 29-SEP-2000; 2000US-0237030P.
PR (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
PI WPI; 2002-372126/40.
DR N-PSDB; AAD35800.
XX New isolated mast cell immunoreceptor signal transducer polypeptide,
PT useful for treating immune disorder involving hyperactivity of B- or T-
PT lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma.
XX Claim 11; Fig 3; 171pp; English.
XX The present invention relates to novel mast cell immunoreceptor signal
CC transducer (MIST) proteins and polynucleotides encoding such proteins.
CC MIST sequences of the invention are useful for preventing, treating or
CC ameliorating a medical condition in mammalian subject. They are useful
CC for treating an immune disorder involving hyperactivity of B- or T-
CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful as
CC targets for therapeutic intervention in immune cell disorders and
CC inflammatory indications, for diagnosis and/or screening of disorders or
CC diseases associated with expression of MIST, for screening for
CC antagonists or inhibitors of the interaction of MIST with cellular
CC signalling components. They are used in assays that detect activation or
CC induction of various B and T-cell-related neoplasms or cancers. Sequences
CC of the invention are also used in gene therapy. The present sequence is
CC human MIST protein

```

XX
SQ Sequence 443 AA;
Query Match 56.8%; Score 1316.5; DB 5; Length 443;
Best Local Similarity 62.1%; Pred. No. 4.6e-106;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;
QY 1 MTSQGNKRTTKKEGFDLRFONVSLKNRSPSLSSAKGRCAVLEPLPDHRRNLGAVPGG 60
DB 16 MNQGNKRTTKEGSNDLKFQNSLPKNSWPRINSATGQYQRMNKPLLDWERNFAAVLDG 75
QY 61 EKCSNNNDYEDPEFQLLKAWPSMKILPARPIQSEYADTRYFQDMMEAPLILPPKASVST 120
DB 76 AKGHSDDYDDPELRMEETWQSIKILPARPIKSEYADTHYFKVAMDTPLDLTRSISI 135
QY 121 EROTDRVMTQLEBVKDPTKDVRSQRFKFKYTKINKTLPPLPPRAITLPPKYQPLPPA 180
DB 136 GQPTWNTQ-TRLERVDKPIKDVRSQNIKGDSVRKNKIPLPPRPPLITLPPKYQPLPP- 193
QY 181 PPESSAYFAPKTFEVRQGRQRSADKFSVLGAEEESHQTKPESSCPSSNQNTQKS 240
DB 194 EPSSRPPLUSQRHTFEVRQMPQSIQLRDLSEVLEAEKVPHNQRPSTHLENQNTQEI 253
QY 241 PPAIASSSYMPGKHSIQARDHTGSMOHCAPQRCQAASHSP--RMLPYENTNSEKDPDK 298
DB 254 PLAISSSFTSNHSVQNRDHRGMPQSPQRCQPPASCSPHENILPYKTSWRPFPR 313
QY 299 PDEKDVQNEWYIGEYSRQAVEDVLAKENKDGTFVLVDCSTKSABPYVLVVFYGNKVYN 358
DB 314 SDRKDVQNEWYIGEYSRQAVEAFMKNKDGSLVDCSTKSKEEPPYVLAVFYENKVYN 373
QY 359 VKIRFLESNQOQFALGTGLRGNEFDSVEDIIIEHYTFPILLIDGDK-AARRKQCVLTQ 417
DB 374 VKIRFERNQOQFALGTGLRGDEKFSVEDIIIEHYKNFPILLIDGDKTGVRKQCHLTQ 433
QY 418 LPLRLALL 425
DB 434 LPLTRHLL 441
RESULT 5
AAB81071
ID AAB81071 standard; protein; 376 AA.
AC AAB81071;
XX
DT 25-JUN-2001 (first entry)
XX
DE Human mast cell-specific immunoreceptor signal transducer.
XX
KW Mast cell; signal transduction; human; allergic disease; MIST;
KW mast cell-specific immunoreceptor signal transducer.
XX
OS Homo sapiens.
XX
PN JP3146204-B1.
XX
PD 12-MAR-2001.
XX
PF 17-SEP-1999; 99JP-00263778.
XX
PR 17-SEP-1999; 99JP-00263778.
XX
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR WPI; 2001-310022/33.
DR N-PSDB; AAF86140.
XX
PT Mast cell-specific signal-transduction molecule, useful for screening
PT therapeutic compounds for treating allergies, is specifically expressed
PT by mouse mast cell.
XX
XX Example 1; Page 10-11; 12pp; Japanese.
PS

```

XX This sequence represents human MIST (mast cell-specific immunoreceptor signal transducer). The invention relates to cDNA encoding a murine mast cell-specific signal transduction protein. Included in the invention are cDNA and protein sequences of the mast cell-specific signal transduction molecule and an expression vector containing the polynucleotide sequence. The coding sequence of the signal transduction protein can be used for screening therapeutic compounds which will be useful for treating allergic diseases

XX SQ Sequence 376 AA;

Query Match 48.9%; Score 1132; DB 4; Length 376;
Best Local Similarity 60.9%; Pred. No. 5e-90;
Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

QY 19 FQNVSLKNRSPSLSSAKGRCAVLEPLPDHRRNLGAVPGGKCSNNNDYEDPEFOLLK 78
DB 1 FQNFSLPKNSWPRINSATGQYQRMNKPLLDWERNFAAVLDGAKGHSDDYDDPELRMEE 60

QY 79 AWPMSKILPARPIQSEYADTRYFQDMMEAPLILPPKASVSTERTQDRVMTQLEEVDPK 138
DB 61 TWQSIKILPARPIKSEYADTHYFKVAMDTPLDLTRSISIGQPTWNTQ-TRLERVDKP 119

QY 139 TFXDVSQRFKFKYTKINKTLPPLPPRAITLPPKYQPLPPAPPESSAYFAPKTPPEV 198
DB 120 ISRDVRSQNIKGDSVRKNKIPLPPRPPLITLPPKYQPLPP-EPSSRPPLSQRHTPPEV 178

QY 199 QRQPRQRSADKFSVLGAEEESHQTKPESSCPSSNQNTQKSPPAIASSSYMPGKHSIQ 258
DB 179 QGMPQSIQLRDLSEVLEAEKVPHNQRPSTHLENQNTQEIPLAISSSFTSNHSVQN 238

QY 259 RDIITGSMOHCAPQRCQAASHSP--RMLPYENTNSEKDPDKDEKDVQNEWYIGEYSR 316
DB 239 RDHRGGMQPCSPQRCQPPASCSPHENILPYKTSWRPFPKRSRDKDVQNEWYIGEYSR 298

QY 317 QAVEDVLAKENKDGTFVLVDCSTKSABPYVLVVFYGNKVYNKIRFLESNQOQFALGTGL 376
DB 299 QAVEEAFMKNKDGSLVDCSTKSKEEPPYVLAVFYENKVYNKIRFLERNQOQFALGTGL 358

QY 377 RGNEMFDSVEDIIIEHY 392
DB 359 RGDEKFSVEDIIIEHY 374

RESULT 6
AAE22610
ID AAE22610 standard; protein; 353 AA.
AC AAE22610;
XX
DT 26-JUL-2002 (first entry)
XX
DE Human MIST splice variant protein from clone #12.
XX
KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
XX
OS Homo sapiens.
XX
PN WO200226986-A2.
XX
PD 04-APR-2002.
XX
PF 28-SEP-2001; 2001WO-US030593.
XX
PR 29-SEP-2000; 2000US-0237030P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX

DR WPI; 2002-372126/40.
XX N-PSDB; AAD35802.
XX New isolated mast cell immunoreceptor signal transducer polypeptide,
PT useful for treating immune disorder involving hyperactivity of B- or T-
PT lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma.
XX Claim 11; Fig 8; 17lpp; English.
XX

The present invention relates to novel mast cell immunoreceptor signal
transducer (MIST) proteins and polynucleotides encoding such proteins.
MIST sequences of the invention are useful for preventing, treating or
ameliorating a medical condition in mammalian subject. They are useful
for treating an immune disorder involving hyperactivity of B- or T-
lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
lymphoma, tumour or thymoma in a mammal. MIST sequences are useful as
targets for therapeutic intervention in immune cell disorders and
inflammatory indications, for diagnosis and/or screening of disorders or
diseases associated with expression of MIST, for screening for
antagonists or inhibitors of the interaction of MIST with cellular
signalling components. They are used in assays that detect activation or
induction of various B and T-cell-related neoplasms or cancers. Sequences
of the invention are also used in gene therapy. The present sequence is
human MIST splice variant protein from clone #12

XX Sequence 353 AA;
SQ

Query Match 47.9%; Score 1109.5; DB 5; Length 353;
Best Local Similarity 64.4%; Pred. No. 4.2e-88;
Matches 226; Conservative 31; Mismatches 89; Indels 5; Gaps 4;

QY 78 KAWPSMKILPARIQSEYADTYFQDMMEAPILLPPKASVSTERQTRVMTQLEVDK 137
Db 3 ETQSIKILPARIQSEYADTYFQDMMEAPILLPPKASVSTERQTRVMTQLEVDK 61

QY 138 PTEKDVRSORFKYTKINKTLPPLPPPAITLPPKYPQPPAPPRESSAYFAPKTFPE 197
Db 62 PISKDVRSQNIKGASVRKNKILFLLPPPRPLITLPPKYPQPPAPPRESSAYFAPKTFPE 120

QY 198 VQGGPRORSADKFSRLVGAEEESHQPKPSSQNQTKSPPAIASSSYMPGKHISIQ 257
Db 121 VQMPQSIURDLSEVLEAKVPHNQKPESTHLENQNTQEIPLAIISSSFTTSHSVQ 180

QY 258 ARDHTGSMQHCAPQRCQAASHSP--RMLPYENTNSKPDTPKDEKDVQNWENYIGEYS 315
Db 181 NRDRGGMQPCSPQRCQPASCSPHENILPKYTSWRPFPKSRDKDVQHNWENYIGEYS 240

QY 316 QRAVEDVLKMKNDGTFVLDGCTSKSAEPYVLVFGNKVYKVRPLESNQOQFALGTG 375
Db 241 QRAVEEAFMKNDGSLFVLDGCTSKSEEPYVLAVFYENKYNKIRFLERNQOQFALGTG 300

QY 376 LRGNEMFDSVEDIEHYTFYPIILLIDGDK-AARRKCYLTQPLPLARLLL 425
Db 301 LRGDEKFDSDVEDIEHYKNFPIILLIDGDKDTGVHRKQCHLTQPLPLTRHLL 351

RESULT 7
ADP65191
XX ADP65191 standard; protein; 533 AA.
AC ADP65191;
XX

12-AUG-2004 (first entry)
XX Human lymphocyte cytosolic protein 2, SH2 domain-containing leukocyte.
DE autoimmune disease; arthritis; gene expression analysis;
XX rheumatoid arthritis; collagen-induced; immunosuppressive; anti-rheumatic;
KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
XX immune; human.

OS Homo sapiens.
XX WO2003072827-A1.
XX 04-SBP-2003.
XX 31-OCT-2002; 2002WO-US035433.
XX 31-OCT-2001; 2001US-0336220P.
XX (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX Hirsch R, Thorton SL;
XX WPI; 2003-712740/67.
XX GENBANK; NP_005556.
XX Diagnosing and analyzing autoimmune disease using gene expression
profiles and microarray technology, useful for diagnosing and treating
rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
gout.
XX Disclosure; Page; 56pp; English.
XX The invention relates to a novel method for diagnosing and analysing
autoimmune disease or arthritides. The method comprises obtaining a
patient sample containing mRNA, analysing gene expression using the mRNA
that results in a gene expression signature of the mRNA, and using that
gene expression signature to diagnose or analyse the autoimmune disease
or arthritides in the patient, where gene expression of at least 60% of
the genes correlates with that of the gene signature. The invention
further comprises: a treatment of rheumatoid arthritis; identification of
genes for targeting in the treatment of rheumatoid arthritis in a mammal
other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
array or gene chip, specific for rheumatoid arthritis; diagnosis or
analyses of autoimmune disease or rheumatoid arthritis; screening the
efficacy of a candidate drug in vitro for the treatment of collagen-
induced arthritis; and reducing the symptoms associated with collagen-
induced arthritis. The compositions of the invention have the following
activities: immunosuppressive, anti-rheumatic, antiarthritic, osteopathic,
antigout, antiinflammatory, dermatological, and immunomodulatory. The
methods and compositions of the present invention are useful for
diagnosing and treating autoimmune disease or arthritides, such as
rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
immune disease caused by an infectious agent. This sequence represents a
protein sequence relating to the genes used in the analysis and treatment
of autoimmune diseases or arthritides. Note: This sequence is not shown
in the specification. It has been supplied in an electronic format from
WIPO.

XX SQ Sequence 533 AA;
Query Match 17.2%; Score 398; DB 7; Length 533;
Best Local Similarity 30.2%; Pred. No. 1.2e-25;
Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;

QY 68 DYEDPEFQLLKAWPSMKILPARIQSE--YATRYFQDMMEAP-----LLLPKKA 116
Db 144 DYEPFNSDEALQN-SILPAKFPNNSMYIDRPPSGKTPQQPPVPPQRPMAALPPPPA 202

QY 117 -----SVSTERQTRVMTQLEVD-----KPT 139
Db 203 GRNHSPLPPQTNEHPPSRNRHKTAKLPAPSIDRSTKPLDRSLAPDFEPFTLGKKPP 262

QY 140 FKD---VRSORFKYTKINKTLPPLPPPAITLPPKYPQPPAPPRESSAYFAPKTFPE 196
Db 263 FSKPSP1PAGRSUJGEHLPIKQKPLPP-----TTERHERSSSLP-----GKKPVP 308

QY 197 EVQGRPQRSADKFSRLVGAEEESHQTKPE-----SSCPSSNQNTQKSPPAIASSSY 250
Db 309 KHGWGPDRE-----NDEDDVHQRPPLQPALPMSSNTFFSRSTKPSPMNPLPSSH 360

QY 251 PGKHSIQARDHTGSMQHCQAQRAAASHSP-----RMLPYENTNSEKPDPTKDEKDVW 305
 Db 361 PGAFS-ESNSFFQSASLPYFSGPSNRPPPIRAEGRNFPLPLPNKPRP-PSPAEENS 418
 QY 306 QNEWYIGEYRQAVEDVLMKENKDGTLVRDCSTKSKAEYPYLVVFGKVNKVFIRFLE 365
 Db 419 NEEWYVSYITRPEAEALRKINQDGTFLVRDSSKTTTNPYVLMVLYKDKVNIQIRYQK 478
 QY 366 SNOQFALGTGLRGNEMFDSVEDIEHYTYPFILLIDGKKAARKQCYLT 415
 Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLLLIDGKNRGS-RYQCTLT 527

RESULT 8
 ADQ19792
 ID ADQ19792 standard; protein; 533 AA.
 XX AC ADQ19792;
 XX DT 26-AUG-2004 (first entry)
 XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 2611.
 XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
 XX OS Homo sapiens.
 XX PN W02004048938-A2.
 XX PD 10-JUN-2004.
 XX PF 26-NOV-2003; 2003WO-US038193.
 XX PR 26-NOV-2002; 2002US-0429739P.
 XX PA (PROT-) PROTEIN DESIGN LABS INC.
 XX PI Aziz N, Ginsburg WM, Zlotnik A;
 XX WPI; 2004-441208/41.
 PT Early detection of soft tissue sarcoma comprises determining expression
 PT of a gene in a first soft tissue sample and a normal soft tissue sample
 PT and comparing the gene expression, also useful in treating soft tissue
 PT sarcoma.
 XX Example 2; SEQ ID NO 2611; 210pp; English.
 CC The invention relates to a novel method for detecting soft tissue sarcoma
 CC which comprises obtaining a first soft tissue sample from an individual
 CC and a normal soft tissue sample from the same or different individual,
 CC determining the expression of a gene in both samples and comparing the
 CC expression of the gene in both soft tissue samples, where a higher level
 CC of protein expression in the first soft tissue sample indicates the
 CC presence of soft tissue sarcoma. The method of the invention has
 CC cyrostatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC protein of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.
 XX Sequence 533 AA;
 SQ Query Match 17.2%; Score 398; DB 8; Length 533;
 Best Local Similarity 30.2%; Pred. No. 1.2e-25;
 Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;
 QY 68 DYEDPEQLLKAWSMKILPARIQSE--YADTRYFQDNMEAP-----LLPPPKA 116
 Db 144 DYEPPPSDNEALQN-SILPAKPFPSNSMYIDRPPSGKTPQPPVPFQRPMAALPPPPA 202
 QY 117 -----SVSTERQTRDVRMTQLEVD-----KPT 139

Db 203 GRNHSPLPPQTWHBPSRSRNHKTAKLPAPSIDRSTKPLDRLSLAPDFREPFTLGKKPP 262
 QY 140 FKD---VRSORFXGFKYTKINKTLPPLPPRPAITLPKKYQPLPAPPPESSAYFAPKPTFP 196
 Db 263 FSDKPSIPAGRSIGEHLPKIQKPLPP-----TTERHERSSPLP-----GKKPVP 308
 QY 197 EVQGRPRQRAKDFSRVLGAEEESHQTKPE-----SSCPSSNQNTQKSPPAIASSYM 250
 Db 309 KHGWGPDRE-----NDEDDVHORPLPQPALPMSSNTFPSSRTKSPMNPPLPSSHM 360
 QY 251 PGKHSIQARDHTGSMQHCQAQRAAASHSP-----RMLPYENTNSEKPDPTKDEKDVW 305
 Db 361 PGAFS-ESNSFFQSASLPYFSGPSNRPPPIRAEGRNFPLPLPNKPRP-PSPAEENS 418
 QY 306 QNEWYIGEYRQAVEDVLMKENKDGTLVRDCSTKSKAEYPYLVVFGKVNKVFIRFLE 365
 Db 419 NEEWYVSYITRPEAEALRKINQDGTFLVRDSSKTTTNPYVLMVLYKDKVNIQIRYQK 478
 QY 366 SNOQFALGTGLRGNEMFDSVEDIEHYTYPFILLIDGKKAARKQCYLT 415
 Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLLLIDGKNRGS-RYQCTLT 527

RESULT 9
 ADP23847
 ID ADP23847 standard; protein; 533 AA.
 XX AC ADP23847;
 XX DT 18-NOV-2004 (first entry)
 XX DE PRO polypeptide SEQ ID NO:1025.
 XX KW PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
 KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
 KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
 XX OS Unidentified.
 XX PN W02004041170-A2.
 XX PD 21-MAY-2004.
 XX PF 30-OCT-2003; 2003WO-US034312.
 XX PR 01-NOV-2002; 2002US-0423394P.
 XX (GETH) GENENTECH INC.
 XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
 XX Wu TD;
 XX WPI; 2004-419628/39.
 XX N-PSDB; ADP23846.
 PT New PRO polypeptides and polynucleotides, useful for treating e.g.
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
 PT renal disease, or demyelinating diseases of the central or peripheral
 PT nervous system.
 XX Claim 7; SEQ ID NO 1025; 2940pp; English.
 CC The invention relates to a novel isolated nucleic acid and the PRO
 CC polypeptide encoded by it. A protein of the invention has
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
 CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its
 CC agonist, antagonist, or antibody that specifically binds to the
 CC polypeptide is useful for treating an immune related disorder such as
 CC systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,
 CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an

CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.
XX
SQ Sequence 533 AA;
Query Match 17.2%; Score 398; DB 8; Length 533;
Best Local Similarity 30.2%; Pred. No. 1.2e-25;
Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;
Qy 68 DYEDPEFQLLKAWPSMKILPARPIQSE--YADTRYFQDMWEAP-----LLLPKA 116
Db 144 DYEPPPSNDEALQN-SILPAKPPFNNSMYIDRPPSGKTPQQPPVPQRPMAALPPPPA 202
Qy 117 -----SVSTERQTRVDMTOLBEVD-----KPT 139
Db 203 GRNHSPLPPQTNHEEPRSRSRHHKTAKLPAPSIDRSTKPLDRSLAPFDREPTLGGKPP 262
Qy 140 FKD---VRSORFKGFKYTKINKTLPPLPPRPAITLKKYQPLPPAPPBESSAYFAPKPTFP 196
Db 263 FSDKPSIPAGSLGELHLPKIQKPLPP-----TTERHERSSPLP-----GKKPPVP 308
Qy 197 EVQGGPQRSKADPSRVLGAEESSHHTKPE-----SSCPSSNQNTQKSPPAIASSYM 250
Db 309 KHGSGPDRRE-----NDEDDVHQRPLOPALLPMSSNTTFFSRSTKSPMNPPLPSHHM 360
Qy 251 PGKHSIQARDHTGSMQHCPAQRQAAASHSP-----BMLPVNTNSKPKPTKDEKDVW 305
Db 361 PGAFS-ESNSFFQSASLPYFSGPSNRPPPIRAEGRNFPPLPNKPRP-PSPAEENSL 418
Qy 306 QNEWYIGYSRQAVEDVLMKENKDGTFVLRDCSTKSKAEPLYVLYFGNKVYNVKIRFLE 365
Db 419 NEWYVSYITPEAEALRKINQDGTFLVRDSSKTTNPVLMVLYKDKVYNQIRYQK 478
Qy 366 SNOQFALGTGLRGNEFDSVEDIEHYTYFPIILLIDGKDKAARKQCVLT 415
Db 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLELLIDGKNRGS-RYQCTLT 527
RESULT 10
ID ADP23845
AC ADP23845 standard; protein; 533 AA.
XX
AC ADP23845;
XX
XX 18-NOV-2004 (first entry)
XX
XX PRO polypeptide SEQ ID NO:1023.
XX
XX PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
XX osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
XX antiaethmatic; hepatotropic; respiratory; gene therapy; immune system.
XX
XX Unidentified.
XX
XX WO2004041170-A2.
XX
XX 21-MAY-2004.
XX
XX

PF 30-OCT-2003; 2003WO-US034312.
XX
XX 01-NOV-2002; 2002US-0423394P.
XX
XX (GETH) GENENTECH INC.
XX
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WJ;
XX Wu TD;
XX
XX WPI; 2004-419628/39.
XX N-PSDB; ADP23844.
XX
XX New PRO polypeptides and polynucleotides, useful for treating e.g.
XX erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
XX renal disease, or demyelinating diseases of the central or peripheral
XX nervous system.
XX
XX Claim 7; SEQ ID NO 1023; 2940pp; English.
XX
XX The invention relates to a novel isolated nucleic acid and the PRO
XX polypeptide encoded by it. A protein of the invention has
XX antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
XX osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
XX antiaethmatic, hepatotropic, and respiratory activity. A polynucleotide
XX of the invention may have a use in gene therapy. The PRO polypeptide, its
XX agonist, antagonist, or antibody that specifically binds to the
XX polypeptide is useful for treating an immune related disorder such as
XX systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
XX juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
XX idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
XX vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
XX thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
XX disease, a demyelinating disease of the central or peripheral nervous
XX system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
XX a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
XX disease, infectious or autoimmune chronic active hepatitis, primary
XX biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
XX inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
XX disease, an autoimmune or immune-mediated skin disease, a bullous skin
XX disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
XX disease, asthma, allergic rhinitis, atopic dermatitis, food
XX hypersensitivity, urticaria, an immunologic disease of the lung,
XX eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
XX pneumonitis, a transplantation associated disease, graft rejection or
XX graft-versus-host disease. The present sequence represents a PRO protein
XX of the invention.
XX
SQ Sequence 533 AA;
Query Match 17.2%; Score 398; DB 8; Length 533;
Best Local Similarity 30.2%; Pred. No. 1.2e-25;
Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;
Qy 68 DYEDPEFQLLKAWPSMKILPARPIQSE--YADTRYFQDMWEAP-----LLLPKA 116
Db 144 DYEPPPSNDEALQN-SILPAKPPFNNSMYIDRPPSGKTPQQPPVPQRPMAALPPPPA 202
Qy 117 -----SVSTERQTRVDMTOLBEVD-----KPT 139
Db 203 GRNHSPLPPQTNHEEPRSRSRHHKTAKLPAPSIDRSTKPLDRSLAPFDREPTLGGKPP 262
Qy 140 FKD---VRSORFKGFKYTKINKTLPPLPPRPAITLKKYQPLPPAPPBESSAYFAPKPTFP 196
Db 263 FSDKPSIPAGSLGELHLPKIQKPLPP-----TTERHERSSPLP-----GKKPPVP 308
Qy 197 EVQGGPQRSKADPSRVLGAEESSHHTKPE-----SSCPSSNQNTQKSPPAIASSYM 250
Db 309 KHGSGPDRRE-----NDEDDVHQRPLOPALLPMSSNTTFFSRSTKSPMNPPLPSHHM 360
Qy 251 PGKHSIQARDHTGSMQHCPAQRQAAASHSP-----BMLPVNTNSKPKPTKDEKDVW 305
Db 361 PGAFS-ESNSFFQSASLPYFSGPSNRPPPIRAEGRNFPPLPNKPRP-PSPAEENSL 418
XX

QY 306 QNEWYIGEYSRQAVEDVLMKENKDGTFVLRDCTSKAEYPVLVVFYGNKVYNKIRFLE 365
 DB 419 NEEWYVSYITRPEAEALRKINQDGTFLVRDSSKTTTNPYVLMVLYKDKVNIQIRYQK 478

QY 366 SNOQFALGTGLRGNEFDSVEDIIEHYTFYPILLIDGKDKAARRKQCYLT 415
 DB 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLIIDGKNRGS-RYQCTLT 527

RESULT 11
 ADT07538 standard; protein; 533 AA.
 XX ADT07538;
 AC ADT07538;
 XX 13-JAN-2005 (first entry)
 DT
 DE Human colon-specific polypeptide (CSP) #55.
 XX Human; colon-specific polypeptide; CSP; metastasis; colon cancer;
 KW Cytostatic.
 KW
 XX Homo sapiens.
 OS
 XX W02004089301-A2.
 PN
 XX 21-OCT-2004.
 PD
 XX 02-APR-2004; 2004WO-US010531.
 PF
 XX 02-APR-2003; 2003US-0460711P.
 PR
 XX (DIAD-) DIADEXUS INC.
 PA
 XX Macina RA, Turner LR, Sun Y;
 PI WPI; 2004-748572/73.
 DR N-PSDB; ADT07466.
 DR
 XX New mammalian or human nucleic acid molecule or polypeptide, useful in
 PT preparing a composition for diagnosing or treating colon cancer.
 PT
 XX Claim 12; SEQ ID NO 125; 596pp; English.
 PS
 XX The invention relates to the isolation of human colon specific nucleic
 CC acids (CSNAs) and the encoding colon-specific polypeptides (CSPs). Also
 CC disclosed is a method for determining the presence of a CSNA in a sample,
 CC a method for diagnosing or monitoring the presence and metastases of
 CC colon cancer in a patient, a method of treating a patient with colon
 CC cancer, a vaccine consisting of the isolated nucleic acid molecule or
 CC polypeptide, and a kit for detecting a risk of cancer or presence of
 CC cancer in a patient. The nucleic acid molecule is useful in preparing a
 CC composition for diagnosing or treating colon cancer. The present sequence
 CC represents a CSP of the invention.
 XX
 SQ Sequence 533 AA;

Query Match 17.2%; Score 398; DB 8; Length 533;
 Best Local Similarity 30.2%; Pred. No. 1.2e-25;
 Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;

QY 68 DYEDPEQLLKAWPSMKILPARIQSE--YADTRYFDQDMEAP-----LLLPKKA 116
 DB 144 DYEPPPSNDDEALQN-SILPAKPPFNNSMYIDRPPSGKTKPQQPPVPPQRPMAALPPPPA 202
 QY 117 -----SVSTEROTDRVMTQLEVD-----KPT 139
 DB 203 GRNHSPLPPQTWHEPSSRNHKTAKLPAPSIDRSTKPPDLRSLAPDFREPFLGKPP 262
 QY 140 FKD---VRSQRFKGYKTKINKTLPPLPPRPAILTKPKYQPLPPAPPESSAYFAPKPTFP 196
 DB 263 FSDKPSIPAGSLGEHLPKIQKPLPP-----TTERHERSSPLP-----GKKPPVP 308

QY 197 EVQGRQRSAKDFSRVLGAEESESHQTKPE-----SSCPSNNQNTQKSPPAISSSYM 250
 DB 309 KHGWGDPDRR-----NDEDDVHORPLPQPALLPMSNTFFSRSTKPSPMNPLPSSHM 360

QY 251 PGKHSIQARDHTGSMOHCFAQRCAAAASHSP-----RMLFYENTNSENKPTKPDKDVW 305
 DB 361 PGAFS-ESNSSFPQSASLPPYFSGPSNRPPIRAEGRNFFPLPLPNKPRP-PSPAEEENSL 418

QY 306 QNEWYIGEYSRQAVEDVLMKENKDGTFVLRDCTSKAEYPVLVVFYGNKVYNKIRFLE 365
 DB 419 NEEWYVSYITRPEAEALRKINQDGTFLVRDSSKTTTNPYVLMVLYKDKVNIQIRYQK 478

QY 366 SNOQFALGTGLRGNEFDSVEDIIEHYTFYPILLIDGKDKAARRKQCYLT 415
 DB 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLIIDGKNRGS-RYQCTLT 527

RESULT 12
 ADP45459
 ID ADP45459 standard; protein; 224 AA.
 XX
 AC ADP45459;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Human colon cancer cell-associated protein SeqID182.
 XX
 KW neoplastic colon cell; cytostatic; vaccine; gene therapy; colon cancer;
 KW immune response; metastasis; human.
 KW
 OS Homo sapiens.
 XX
 PN W02004050900-A2.
 XX
 PD 17-JUN-2004.
 XX
 XX 04-DEC-2003; 2003WO-US040131.
 PF
 XX 04-DEC-2002; 2002US-0431143P.
 PR 04-DEC-2002; 2002US-0431206P.
 XX
 PA (DIAD-) DIADEXUS INC.
 XX
 XX Macina RA, Turner LR, Sun Y, Rodriguez M;
 PI WPI; 2004-480629/45.
 DR N-PSDB; ADP45333.
 DR
 XX Nucleic acid molecules isolated from normal and neoplastic colon cells,
 CC useful for treating colon cancer and for identifying, diagnosing,
 CC monitoring, staging, and imaging colon cancer and/or non-cancerous
 CC disease states in colon.
 PT
 XX Claim 12; SEQ ID NO 182; 564pp; English.
 PS
 XX This invention relates to novel nucleic acid molecules isolated from
 CC normal and neoplastic colon cells and the proteins encoded by them. The
 CC invention may be useful for the production of compounds with a cytostatic
 CC activity or for the development of a vaccine or gene therapy. The
 CC invention may be useful for developing a treatment for a patient with
 CC colon cancer, where the administration induces an immune response against
 CC the colon cancer cell expressing the novel nucleic acids or proteins. The
 CC invention may also be useful for diagnosing or monitoring the presence
 CC and metastases of colon cancer in a patient. The present sequence is that
 CC of a colon cancer cell-associated protein of the invention.
 XX
 SQ Sequence 224 AA;

Query Match 14.6%; Score 337; DB 8; Length 224;
 Best Local Similarity 40.0%; Pred. No. 7.7e-21;
 Matches 78; Conservative 33; Mismatches 76; Indels 8; Gaps 4;

QY 226 PESSCPSNNQNTQKSPPAISSSYMFGKHSIQARDHTGSMOHCFAQRCAAAASHSP----- 281


```

Db      |||  ::::: |||  |||: |||  | ::|  | ::|  | ::|  |
27 PMSNTFSPSRSTKPSMNPLPSSHMPGAFS-ESNSFPQSASLPPYFSQGPSNRPPIRAE 85

Qy      -RMLPYENTNSEKPDPTKPEKQVWQNEWYIGEYSRQAVEDVLMKENKDGTFLLVRDCSTK 340
      |||  | ::|  | ::|  | ::|  | ::|  | ::|  | ::|  |
86 GRNFPPLPNKPRP-PSPAEENSLSNEWYVYITRPEAEAAARKINQDGTFLVRDSSKK 144

Qy      SKAEPYVLVVPYGNKVNVKIRFLESNQOQFALGTGLRGNEMFDSVEDIIHHTYTFPILLI 400
      : |||: | ::|: | ::|: | ::|: | ::|: | ::|: | ::|: |
145 TTNPFYVLMVLYKQVNIQIRYOKESQVYLLGTGLRKGEDFLSVSDIIDYFRKMPLLLI 204

Qy      401 DGKDKAARRKQCYLT 415
      |||: : | |||
205 DGNRGS-RYQCTLT 218

Db**

RESULT 13
ADP45461
ID ADP45461 standard; protein; 225 AA.
XX
AC ADP45461;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human colon cancer cell-associated protein SeqID184.
XX
KW neoplastic colon cell; cytostatic; vaccine; gene therapy; colon cancer;
KW immune response; metastasis; human.
XX
OS Homo sapiens.
XX
PN WO2004050900-A2.
XX
PD 17-JUN-2004.
XX
PF 04-DEC-2003; 2003WO-US040131.
XX
PR 04-DEC-2002; 2002US-0431143P.
PR 04-DEC-2002; 2002US-0431206P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Macina RA, Turner LR, Sun Y, Rodriguez M;
XX
DR WPI; 2004-480629/45.
DR N-PSDB; ADP45334.
XX
PT Nucleic acid molecules isolated from normal and neoplastic colon cells,
PT useful for treating colon cancer and for identifying, diagnosing,
PT monitoring, staging, and imaging colon cancer and/or non-cancerous
PT disease states in colon.
XX
PS Claim 12; SEQ ID NO 184; 564pp; English.
XX
CC This invention relates to novel nucleic acid molecules isolated from
CC normal and neoplastic colon cells and the proteins encoded by them. The
CC invention may be useful for the production of compounds with a cytostatic
CC activity or for the development of a vaccine or gene therapy. The
CC invention may be useful for developing a treatment for a patient with
CC colon cancer, where the administration induces an immune response against
CC the colon cancer cell expressing the novel nucleic acids or proteins. The
CC invention may also be useful for diagnosing or monitoring the presence
CC and metastases of colon cancer in a patient. The present sequence is that
CC of a colon cancer cell-associated protein of the invention.
XX
SQ Sequence 225 AA;

Query Match 14.6%; Score 337; DB 8; Length 225;
Best Local Similarity 40.0%; Pred. NO. 7.7e-21;
Matches 78; Conservative 33; Mismatches 76; Indels 8; Gaps 4;

Qy      |||  ::::: |||  |||: |||  | ::|  | ::|  | ::|  |
226 PESSCPSSNQTOKSPPAIASSSYMPGKHSTQARDHTGSMQHCQAQRCAAAASHSP---- 281
      |||  | ::|: | ::|: | ::|: | ::|: | ::|: | ::|: |
33 PMSNTFSPSRSTKPSMNPLPSSHMPGAFS-ESNSFPQSASLPPYFSQGPSNRPPIRAE 91

```

```

Db      28 PMSNTFSPSRSTKPSMNPLPSSHMPGAFS-ESNSFPQSASLPPYFSQGPSNRPPIRAE 86

Qy      282 -RMLPYENTNSEKPDPTKPEKQVWQNEWYIGEYSRQAVEDVLMKENKDGTFLLVRDCSTK 340
      |||  | ::|  | ::|  | ::|  | ::|  | ::|  | ::|  |
87 GRNFPPLPNKPRP-PSPAEENSLSNEWYVYITRPEAEAAARKINQDGTFLVRDSSKK 145

Qy      341 SKAEPYVLVVPYGNKVNVKIRFLESNQOQFALGTGLRGNEMFDSVEDIIHHTYTFPILLI 400
      : |||: | ::|: | ::|: | ::|: | ::|: | ::|: | ::|: |
146 TTNPFYVLMVLYKQVNIQIRYOKESQVYLLGTGLRKGEDFLSVSDIIDYFRKMPLLLI 205

Qy      401 DGKDKAARRKQCYLT 415
      |||: : | |||
206 DGNRGS-RYQCTLT 219

Db

RESULT 14
ADP45460
ID ADP45460 standard; protein; 230 AA.
XX
AC ADP45460;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human colon cancer cell-associated protein SeqID183.
XX
KW neoplastic colon cell; cytostatic; vaccine; gene therapy; colon cancer;
KW immune response; metastasis; human.
XX
OS Homo sapiens.
XX
PN WO2004050900-A2.
XX
PD 17-JUN-2004.
XX
PF 04-DEC-2003; 2003WO-US040131.
XX
PR 04-DEC-2002; 2002US-0431143P.
PR 04-DEC-2002; 2002US-0431206P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Macina RA, Turner LR, Sun Y, Rodriguez M;
XX
DR WPI; 2004-480629/45.
DR N-PSDB; ADP45334.
XX
PT Nucleic acid molecules isolated from normal and neoplastic colon cells,
PT useful for treating colon cancer and for identifying, diagnosing,
PT monitoring, staging, and imaging colon cancer and/or non-cancerous
PT disease states in colon.
XX
PS Claim 12; SEQ ID NO 183; 564pp; English.
XX
CC This invention relates to novel nucleic acid molecules isolated from
CC normal and neoplastic colon cells and the proteins encoded by them. The
CC invention may be useful for the production of compounds with a cytostatic
CC activity or for the development of a vaccine or gene therapy. The
CC invention may be useful for developing a treatment for a patient with
CC colon cancer, where the administration induces an immune response against
CC the colon cancer cell expressing the novel nucleic acids or proteins. The
CC invention may also be useful for diagnosing or monitoring the presence
CC and metastases of colon cancer in a patient. The present sequence is that
CC of a colon cancer cell-associated protein of the invention.
XX
SQ Sequence 230 AA;

Query Match 14.6%; Score 337; DB 8; Length 230;
Best Local Similarity 40.0%; Pred. NO. 8e-21;
Matches 78; Conservative 33; Mismatches 76; Indels 8; Gaps 4;

Qy      |||  ::::: |||  |||: |||  | ::|  | ::|  | ::|  |
226 PESSCPSSNQTOKSPPAIASSSYMPGKHSTQARDHTGSMQHCQAQRCAAAASHSP---- 281
      |||  | ::|: | ::|: | ::|: | ::|: | ::|: | ::|: |
33 PMSNTFSPSRSTKPSMNPLPSSHMPGAFS-ESNSFPQSASLPPYFSQGPSNRPPIRAE 91

```

QY 282 -RMLPYENTNSEKDPDKDEKDVQWQNEWYIGESVROAEDVLMKENKCGTFLVRCSTK 340
Db 92 GRFPPLPLPNKPRP-PSPABEENSLEEWVSVYITRPEAEALRKINQDGTFLVROSSKK 150
QY 341 SKAEPYVLYVYFGKVNKIRFLESNQFQALGTGLRGNEMFDSVEDIIBHYTYFPILLI 400
Db 151 TTINPYVLMVLYKDKVNIQIRYKESQVLLGTGLRGKEDFLSVSDIIDIYFRKMPILLI 210
QY 401 DGKDKAARRKQCVILT 415
Db 211 DGNRGS-RYQCTLT 224

RESULT 15
AAB93406
ID AAB93406 standard; protein; 456 AA.
XX AAB93406;
XX
DT 26-JUN-2001 (first entry)
XX
XX Human protein sequence SEQ ID NO:12602.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.
XX
PN EP1074617-A2.
XX
XX 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
PS Claim 8; SEQ ID NO 12602; 2537pp + Sequence Listing; English.
XX
XX The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dr primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAB03166 to AAB13628 and
CC AAB13633 to AAB18742 represent human cDNA sequences; AAB92446 to AAB95893

CC represent human amino acid sequences; and AAB13629 to AAB13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 456 AA;
Query Match 14.1%; Score 326; DB 4; Length 456;
Best Local Similarity 24.6%; Pred. No. 1.9e-19;
Matches 99; Conservative 50; Mismatches 141; Indels 112; Gaps 9;
QY 64 NSNNYEDPEFQLLKAWPSMKILPARPIQSEYADTRYFQDMMEAP---LLLPPKASVST 120
Db 91 NADDSYEPPEVE---QETRPVHPALPFARGEYIDNRSSQ--RHSPPFSKTLSPKSPWPS 144
QY 121 ERQTRDVRMTQLEEVDPKPTFKVRSQRFKGYKTKINKTLPPLPPPAITLTKYQPLPPA 180
Db 145 EKARLTSTLPALTALQKPV-----PPKPKGLLEDEADYVVPV 182
QY 181 PPESSAYFAPKPTFPFVQVQRPQRSKQFSLVGAEEESHHTKPESSCPSSNQNT--- 237
Db 183 EDENYIIHTESSPPPEKAPM-----VNRSTKNSSTPASPPGTASG 226
QY 238 -----QKSPPAIASS-----SYMPKGHSIQARDHTGSMQ 266
Db 227 RNSGAWETKSPPPAAPSPPLPRACKKPTTLPKTPVASQQNASVCGEKPIPAERHRGSSH 286
QY 267 HCPAQRCQAAASHSPMLPYENTNSKDPD-----TKPDEK 302
Db 287 -----RQEAQSPVFPFPAQKIQHOKPIPLPRFTEGNGNPTVDGPLPFSFSSNSTISEQEA 339
QY 303 DVWQNEWYIGESVROAEDVLMKENKDGTFVREDCTSKAEPYVLYVYFGKVNKIR 362
Db 340 GVLCKPWYAGACDRKGAEEALHRSNKGSLIRKSSGHDSDKQPYTTLVFPFKRVYNIPIVR 399
QY 363 FLESNQCFALGTGLRGNEMFDSVEDIIEHYTYFPILLIDGKD 404
Db 400 FIETKQYALGRKKNSEYFGSVAEIIRNHQHSPLVLIDSQN 441

Search completed: March 22, 2006, 01:31:56
Job time : 192 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2006, 01:32:14 ; Search time 41 Seconds
(without alignments)
1020.836 Million cell updates/sec

Title: US-10-717-619-2
Perfect score: 2316
Sequence: 1 MTSQGNKRTTKGEGDLRFQ.....QPLPLARLLLTQYSSQALHE 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412.5	17.8	533	2 B56110	tyrosine phosphor
2	398	17.2	533	2 A56110	tyrosine phosphor
3	154.5	6.7	663	1 TVMVR	protein-tyrosine k
4	146	6.3	728	2 H59435	phosphoinositide-3
5	142.5	6.2	723	2 B38749	3-phosphatidylinos
6	138.5	6.0	443	2 T27877	hypothetical prote
7	134.5	5.8	480	2 JC7552	Shb-like adapter p
8	134	5.8	1599	2 T15854	hypothetical prote
9	133	5.7	675	2 S60612	protein-tyrosine k
10	132.5	5.7	592	1 LLBY	actin-binding prot
11	132	5.7	1603	2 S17983	gene posterior sex
12	130.5	5.6	724	2 A38749	3-phosphatidylinos
13	130	5.6	559	2 T49444	SH3 binding protei
14	130	5.6	724	2 A38747	phosphatidylinosit
15	129	5.6	751	2 S68957	adhesive plaque pr
16	128.5	5.5	3421	1 W2B866	367K tegument prot
17	128	5.5	706	2 A45990	junctional sarcopt
18	128	5.5	724	2 A38748	3-phosphatidylinos
19	127	5.5	1385	2 T21706	hypothetical prote
20	125.5	5.4	1029	2 T30351	mucin-like protein
21	125	5.4	461	2 A57463	p55PIK - mouse
22	125	5.4	962	2 S58107	hypothetical WD-40
23	124.5	5.4	473	2 S25776	transforming prote
24	124.5	5.4	585	2 A46209	protein-tyrosine-p
25	124.5	5.4	593	1 JN0805	protein-tyrosine-p
26	124.5	5.4	597	1 A35193	protein-tyrosine-p
27	123.5	5.3	841	2 A43254	protein-tyrosine-p
28	123.5	5.3	1317	2 T03748	apoptosis associat
29	122.5	5.3	596	2 I38228	Shb protein - huma

30	122.5	5.3	875	2 S23760	polyphenolic adhes
31	121	5.2	485	2 T27147	hypothetical prote
32	120.5	5.2	431	2 T01557	hypothetical prote
33	120.5	5.2	1290	2 A36466	1-phosphatidylinos
34	120.5	5.2	1792	2 A57075	tensin - chicken (
35	120	5.2	559	2 B44265	ENL (translocation
36	120	5.2	981	1 FOMVGM	gag-abl polyprotei
37	119.5	5.2	509	1 TVHAST	protein-tyrosine k
38	119.5	5.2	595	1 A55651	protein-tyrosine-p
39	119.5	5.2	1052	2 A82127	ribonuclease E VC2
40	119.5	5.2	1733	2 S27939	tensin - chicken
41	119.5	5.2	1744	2 A54970	tensin, cardiac mu
42	119	5.1	720	2 T26819	hypothetical prote
43	119	5.1	2187	2 T30826	nascant polypeptid
44	118.5	5.1	440	2 JC7807	Wiskott-Aldrich sy
45	118.5	5.1	657	2 A39526	probable protein k

ALIGNMENTS

RESULT 1

B56110
tyrosine phosphoprotein SLP-76 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 09-Jul-2004
C;Accession: B56110
R;Jackman, J.K.; Motto, D.G.; Sun, Q.; Tanemoto, M.; Turck, C.W.; Peltz, G.A.; Koretzky, J. Biol. Chem. 270, 7029-7032, 1995
A;Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with G G
A;Reference number: A56110; MUID:95221345; PMID:7706237
A;Accession: B56110
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-533 <JAC>
A;Cross-references: UNIPROT:Q60787; UNIPARC:UPI00000284C7; GB:U20159; NID:g806767; PIDN:
C;Keywords: phosphoprotein
F;422-520/Domain: SH2 homology <SH2>

Query Match	17.8%	Score	412.5	DB 2	Length	533
Best Local Similarity	29.7%	Pred. No.	2.5e-20			
Matches	148	Conservative	51	Mismatches	175	Indels 125; Gaps 20;
Qy	14	FGDLRFQNVLL-----KRSWPSLSAKGRCRAVLEPLDPRHRLAG-----	56			
Db	57	FXKLRLPPLSKLQDINKNEERSIFTRKQPIRFLFEETESHEEDDGGWSSFEDDYESPN 116				
Qy	57	--VPGGEKCSNNYEDP--EFQLL-----KAWPS-----MKTLPARPIQ--ESEY 96				
Db	117	DDPDGE---DDGDYSPNEEQALVDDAADYEPFPPSPNNEALQSSILPNSEFHTNSMY 173				
Qy	97	ADTRYFQDMMEAPLLLP--PKASV-----	118			
Db	174	IDRPPTGKVSQQPVPLRPKLPALPLPTGHNHSPSPHPNHEEPSRSGNNKTAKLPAP 233				
Qy	119	STERQTRDVRMTOLEVD-----KPTFKDVRSRFGKFKYTKINKTLPPPRPAITL 170				
Db	234	SIDRSTKPLDRLSLALDRBPFILGKKPFPSDKPSAPLGRHLPKIQKPLP---PAMDR 290				
Qy	171	PKYQPLPAPPRESSAYFAPKPTFFEVQRPQRSADFSRVLGAEESSHQTKPSSC 230				
Db	291	HERNERLGPVTR-----KPSVPRHGRGPRRE-----NDEDDVHQRLPQPQSL 334				
Qy	231	PSSNQNTQKSPPAIAGS-----SYMPGKHSIQARDHTGSMQHC--PAQRCAAAASHP- 281				
Db	335	PSMSSNTFFRSVQPSKNTFFLAHWGAFS--ESNIGFQQSASLPYPSQGPGNRPL 391				
Qy	282	-----RMLPYENTSEKPDPTKPEKOV-WQNEWYIGYSQAVEDVLKMNKGDTFLVRD 336				
Db	392	RSEGNLPLVPVN--RPQPPSGEETPLDEEYVSYITRPEAAALRKINQDGTFLVRD 449				
Qy	337	CSYKSAEPYVLVVFYGNKYVKIRFLESNQFPALGTGLRGNEFMFDSVDITEHYTFP 396				

QY 223 QTKPESSCPSSNQNTQ-KSPPAIASSYMPGKHSIQARDHTGSMQHCQPAORCOAAASHSP 261
||| : : : : :
Db 325 -EGPEXSLSPGREGKRLPRLSAGN-----PKSAKP 356
QY 282 RMLPYENTNSEKDPDKPDQKQVQNEWYIGEYSRQAVEDVLKMKENKDGTFVLVDCSTKS 341
: : : : :
Db 357 LSNEPSPGLGWTDPALPLENOV-----WYIGALSRTDAEN-LRLCKEASLYVRNSETSK 411
: : : : :
QY 342 KABPYVLVWFYGNKVNKVFIRFLESNQFALGTGLRGNEFMFSDVEDIIEHY 392
: : : : :
Db 412 --NDFSLSLKSSQGFQHEMKLSRTKNN-KYVLG---QNSPPFSSVPEIVHHY 456
RESULT 8
T15854
hypothetical protein C56C10.12 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T15854
R:Fulton, L.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid C56C10.
A:Reference number: Z18417
A:Accession: T15854
A:Status: preliminary; translated from GR/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1599 <FUL>
A:Cross-references: UNIPROT:Q18892; UNIPARC:UPI000007A944; EMBL:U29488; NID:g868238; PII
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: C56C10.12
A:Gene: C56C10.12
A:Introns: 38/3; 59/2; 92/3; 178/3; 380/1; 424/3; 554/1; 573/3; 593/2; 697/1; 738/3; 980
C:Superfamily: Caenorhabditis elegans hypothetical protein C56C10.12
Query Match 5.8%; Score 134; DB 2; Length 1599;
Best Local Similarity 23.7%; Pred. No. 0.54; Indels 64; Gaps 13;
Matches 56; Conservative 34; Mismatches 82; Mismatches 82; Indels 64; Gaps 13;
QY 85 ILPARPIQSEYADTRYFQDMMEAP-----LLLP-----PKASVSTERQTRDVRM 129
: : : : :
Db 5 VRPKKP-KHREARQSDYTPPIPRPAIVNGLRLPAAPSHTIEDLPERSAPPEPQDISL 63
QY 130 TQ-----LEEVDKTFQDVRSQKFGKYKINTPTLPPRPATILPKKYQPLPP----- 179
: : : : :
Db 64 TEDSLKREASEP-LKDVRSPPVR-----PAPPP-PRVSQREAPPPIPPRSMIFP 111
QY 180 -----APPESSAYFAPK-----PTFPVQGRQGRSQAQFSRVL-----GAEEES 220
: : : : :
Db 112 RSTSMVAESRKESTTAVAPKRSVAVASYPAVPELAELPSYTDALQHPQVTPSINGLQHS 171
QY 221 HHQT-----KPSSCPSSNQNTQ--KSPPAIAS--SSYMPGKHSIQARDHTGSMQHC 269
||| : : : : :
Db 172 HSATAIPEKTRFSAPEVEREVRGEAPPMPYPSIKTYERNEHGLMTEENLVTFYHNP 227
RESULT 9
S60612
protein-tyrosine kinase (EC 2.7.1.112) Bmx - human
C:Species: Homo sapiens (man)
C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 05-Oct-2004
C:Accession: S60612
R:Tamagnone, L.; Lahtinen, I.; Mustonen, T.; Virtaneva, K.; Francis, F.; Muscatelli, F.;
Oncogene 9, 3683-3688, 1994
A:Title: BMX, a novel nonreceptor tyrosine kinase gene of the BTK/ITK/TEC/TXK family loc
A:Reference number: S60612; MUID:95060827; PMID:7970727
A:Accession: S60612
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-675 <TAM>
A:Cross-references: UNIPROT:PS1813; UNIPARC:UPI0000000DFF; EMBL:X83107; NID:g951234; PII
C:Genetics:
A:Gene: Bmx

C:Superfamily: Tyrosine-protein kinase, TEC type; pleckstrin repeat homology; protein ki
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F;3-109/Domain: pleckstrin repeat homology <PLK>
F;296-392/Domain: SH2 homology <SH2>
F;415-673/Domain: protein kinase homology <KIN>
F;423-431/Region: protein kinase ATP-binding motif
Query Match 5.7%; Score 133; DB 2; Length 675;
Best Local Similarity 24.0%; Pred. No. 0.21;
Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;
QY 166 PATTLPEKKYQPLPAPPPESSAYFAFKPTFPE-VQRGPR-----ORSAKDFSRLGA--- 216
||| : : : : :
Db 141 PGCTLWEAYANLTAVNEKHV-----PTFPDRVLKIPRAVPVLKWDAPSSSTTLAQYDN 196
QY 217 EESHHQTKPESSCPSSNQNTQKSPPAIAS-----SYMPG-----KHSTQAR 259
: : : : :
Db 197 ESKKNYGSQPPSSSTSLAQYDSNKKIYGSQPNFNQYIPREDFPDWMQVRLKSSSSSE 256
QY 260 DHTGSMQHCQPAORCOAAASHSPRLPYENTNSEKPDPTPKDEKDVQNEWYIGYGRQAV 319
: : : : :
Db 257 DVASSNQK-----ERNVNHTTSKISWEFPSSSSE-----EENLDDYDWFAGNISRSQS 306
QY 320 EDVLKMKENKDGTFVLVDCSTKSKAEPYVLVVF---YGNKVNKVFIRFLESNQFALGTGL 376
||| : : : : :
Db 307 EQLLRQKQEGAFPMVRN---SSQVGYTVLSLFSKAVNDKKGTVKHVHVHTNAENKL--YL 361
QY 377 RGNEMFSDVEDIIEHYTY 394
||| : : : : :
Db 362 AENYCFDSIPKLIHYHQH 379
RESULT 10
LLBY
actin-binding protein ABP1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YCR088w
C:Species: Saccharomyces cerevisiae
C:Date: 31-Mar-1991 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: S19503; S19767; S07608
R:Dusterhoft, A.; Erdmann, D.; Hegemann, J.; Philippaen, P.; Schweitzer, B.; Spiegelberg,
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19351
A:Accession: S19503
A:Molecule type: DNA
A:Residues: 1-476 <DUS>
A:Cross-references: UNIPROT:PI15891; UNIPARC:UPI0000173FC9; EMBL:X59720; GSPDB:GN00003; M1
R:Frontali, L.; Grisanti, P.
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19504
A:Accession: S19767
A:Molecule type: DNA
A:Residues: 14-592 <FRO>
A:Cross-references: UNIPARC:UPI0000173FCA; EMBL:X59720; GSPDB:GN00003; MIPS:YCR088w
R:Drubin, D.G.; Mulholland, J.; Zhu, Z.; Botstein, D.
Nature 343, 288-290, 1990
A:Title: Homology of a yeast actin-binding protein to signal transduction proteins and my
A:Reference number: S07608; MUID:90136906; PMID:2405279
A:Accession: S07608
A:Molecule type: DNA
A:Residues: 1-57, 'S', '59-311, 'I', 313-592 <DRU>
A:Cross-references: UNIPARC:UPI0000168AD8; EMBL:X51780; NID:g33321; PIDN:CAA36075.1; PID:9
C:Genetics:
A:Gene: SGD:ABP1; MIPS:YCR088w
A:Cross-references: SGD:S0000684; MIPS:YCR088w
A:Map position: 3R
C:Superfamily: actin-binding protein ABP1; SH3 homology
C:Keywords: actin binding; duplication
F;158-180,378-401/Region: duplication
F;539-588/Domain: SH3 homology <SH3>
Query Match 5.7%; Score 132.5; DB 1; Length 592;
Best Local Similarity 21.4%; Pred. No. 0.19;
Matches 54; Conservative 44; Mismatches 99; Indels 55; Gaps 11;

RESULT 12
A38749
3-phosphatidylinositol kinase (EC 2.7.1.-) 85K chain alpha - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 09-Jul-2004
C;Accession: A38749

Query Match	5.6%	Score 130.5;	DB 2;	Length 724;
Best Local Similarity	20.2%;	Pred. No. 0.34;		
Matches	72;	Conservative	43;	Mismatches 114;
				Indels 127;
				Gaps 15
Qy	148	FKGFKYTKINKTLP	PPRPBATTLP	PKKYQPLPPAP--PESSA-----VFAPK 192
		:::::	:::::	:::::
Db	69	PPGTVEYVIGKKISPTPK---	PPRPRLPVAPGPKSTEAD	SEQOASTLPDLAQFAPP 125
		:::::	:::::	:::::
Qy	193	PTPEVQGRQRSAKDFSRV	LGAEESSHHQTKPSS-----	CPSSNQNTQ----- 238

```

RESULT 13
I49444
SH3 binding protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I49444
C:Ren, R.; Mayer, B.J.; Cicchetti, P.; Baltimore, D.
Science 259, 1157-1161, 1993
A:Title: Identification of a ten-amino acid proline-rich SH3 binding site.
A:Reference number: I49444; MUID:93174278; PMID:8438166

```


A;Accession: I49444
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-559 <RES>
A;Cross-references: UNIPROT:Q06649; UNIPARC:UPI0000027A0E; GB:L14543; NID:g293267; PIDN:
F;25-128/Domain: pleckstrin repeat homology <PLK>
F;201-210/Region: proline-rich SH3 binding

Query Match 5.6%; Score 130; DB 2; Length 559;
Best Local Similarity 21.4%; Pred. No. 0.27;
Matches 87; Conservative 42; Mismatches 145; Indels 132; Gaps 19;

QY 86 LPARPIQESY---ADTRYFQDMMEAPLLI-----PPKASVSTERQTRVMT 130
DB LSSYPMDNEDYEHEDSDSYLSPDPSGPMKLEDAITYPPAYPPPPVPR-----213
QY 131 QLBVDKPTFKDVRQRFKFKYTKINKTP-LPPRPAILTP-----KKY 174
DB LSSYPMDNEDYEHEDSDSYLSPDPSGPMKLEDAITYPPAYPPPPVPR-----213
QY 214 -----KPAFSDLP---RAHSFTSKSPSPLLPPPPPKGLPDTGSAPEDAKDALGLRRV 263
QY 175 QP---LPPAPPESYAFAPKPTFPEVQRPORS-----A 207
DB EPLGRVPATPRRMSDPPMSNVPTPNLRKHPCFRDSVNPGLPWTGHTGTSVSSSTTWA 323
QY 208 KDFSRVLGAEEESHOTK--PESSCPSSNONTOK-----SPPAIASSVMPGKHSIQ 257
DB VATSRNCDKLKSFHLSRGSGPTSEPPVPANKFKFLKIAEPPSPREAAKFAVPV---PVA 380
QY 258 ARDHTGSMQHCAPQRCQAASHSPRLPYENTNSEKPD-----PTKPD--- 300
DB PRPPVQKMP-MPEATVRPAVLPRPENTPLPHLQRSPPDGQSGFRGFEKARQPSQADTGE 439
QY 301 -----EKDVQNEWYIGYSRQAVEDVLMK-----ENKDGTLVRDCSTKSKAEYVL 348
DB EDSDEYKVPLENSVFVNTTESCEVERLFKATDPRGEQDGLYCIIRNSTKSGK---VL 496
QY 349 VVF-YGNKYVNVKIRFLESNQOQFALGTGLRGNEFDSVEDIIEHY 392
DB VVWDESNKVRNRYI--FEKDSKFY----LEGEVLFASVGSVMVEHY 536

RESULT 14
A38747
Phosphatidylinositol 3-kinase (EC 2.7.1.-) 85K chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 09-Jul-2004
C;Accession: A38747
R;Escobedo, J.A.; Navankasattusas, S.; Kavanagh, W.M.; Milfay, V.A.; William
Cell 65, 75-82, 1991
A;Title: cDNA cloning of a novel 85 kd protein that has SH2 domains and regulates bindin
A;Reference number: A38747; MUID:91191564; PMID:1849460
A;Accession: A38747
A;Molecule type: mRNA
A;Residues: 1-724 <ESC>
A;Cross-references: UNIPROT:P70304; UNIPROT:Q8K3B3; UNIPARC:UPI000017C725; GB:M50651
C;Comment: This protein binds a phosphotyrosine-containing sequence of ligand-activated
phosphatidylinositol at position 3 of the inositol ring.
C;Keywords: phosphotransferase
F;333-428/Domain: SH2 homology <SH2A>
F;624-718/Domain: SH2 homology <SH2>

Query Match 5.6%; Score 130; DB 2; Length 724;
Best Local Similarity 22.0%; Pred. No. 0.37;
Matches 89; Conservative 46; Mismatches 147; Indels 122; Gaps 19;

QY 38 GRCAVLEPLPDHR---RNLAGVGGKCNNDYED-PEFQLLKWPMSMKILPARIQES 94
DB GRKX-ISPPTFKRPRPLPVARGSKTEADTEQQALPLPDLAEQFAPDPVAPLLIKLL 136
QY 95 FYADTRYFQDMBAPILLPPKASVSTERQTR-----DVRMTQLEEVDKPTFKDVRQRFK 149
DB EAIEKK-----GLECSTLYRTQSS-SNPAELRQLDCCAAASVDLMDIVHVLADA---FK 187

QY 150 GFXYTKINKTFLPP-----PRPAITLPKKYQPLPPAPPE-ESSAYPAK 192
DB --RYLADLPNPVIVAVYNNMMSLAQELQSPEDCIQLLKKILRLPNIPHOCWTLQLDLK 245
QY 193 PTFPEVQRPQRQSAKD-----FSRVL-----GAEESHHTQTKPESSCPSSNQNT 237
DB 246 HFFK-----LSQASSKNLLNARVLSEIFSPVLPFRPPAASDNTLHLIKATEILISTEWNE 300
QY 238 QKSPPAIASSSYNMPGKHSIQARDHTGSMQHCAPQRCQAASHSPRLPYENTNSEKPDPT 297
DB 301 RQAPALP-----PPKPP 312
QY 298 KP-----DEKDVQNEWYIGYSRQAVEDVLMKENKDGTLVRDCSTKSKAEYVL 348
DB 313 KPTTVANNMNNMNSQDAEWYMGDISREEVNEKL-RDTADGTLVRDASTKMHGD-YTL 370
QY 349 VVFYGNKYVNVKIRFLESNQOQFALGTGLRGNEFDSVEDIIEHY 392
DB 371 TLRKGGN-NKLKIFHRDQKGYGSDPL---TFNSVVELINHY 408

RESULT 15
S68957
adhesive plaque protein precursor - Mediterranean mussel
C;Species: Mytilus galloprovincialis (Mediterranean mussel)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S68957
R;Inoue, K.; Takeuchi, Y.; Miki, D.; Odo, S.; Harayama, S.; Waite, J.H.
Submitted to the EMBL Data Library, August 1995
A;Description: Cloning, sequencing and sites of expression of genes for the hydroxyarginin
A;Reference number: S68957
A;Accession: S68957
A;Molecule type: mRNA
A;Residues: 1-751 <INO>
A;Cross-references: UNIPROT:Q27409; UNIPARC:UPI000012AB79; EMBL:D63778; NID:g961463; PIDN:
C;Keywords: hydroxyproline
F;1-24/Domain: signal sequence #status predicted <SIG>
F;45-751/Product: adhesive plaque protein #status predicted <MAT>
F;123,127,133,173,177,183,227,267,283,293,297,303,307,313,323,327,333,337,343,347,367,373;
33/Modified site: 4-hydroxyproline (Pro) #status predicted
F;125,129,135,139,175,179,185,189,225,229,265,269,285,289,295,299,305,309,315,319,325,329;
79/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status predicted
F;126,136,176,186,226,266,286,296,306,316,326,336,346,366,376,386,406,416,426,436,446,456;
ans-2,3-cis-3,4-dihydroxyproline (Pro) #status predicted
F;485,489,495,499,505,509,515,519,525,529,535,539,545,549,555,559,565,569,575,579,585,589;
dihydroxyphenylalanine (Tyr) #status predicted
F;537,543,547,553,567,573,577,583,587,593,597,603,607,613,617,623,627,633,637,643,647,653;

Query Match 5.6%; Score 129; DB 2; Length 751;
Best Local Similarity 23.4%; Pred. No. 0.45;
Matches 55; Conservative 34; Mismatches 108; Indels 38; Gaps 9;

QY 84 KILPARPIQESYADTRY---FQDMMEAPLLLPKASVSTERQTRDVRMTQLEVDKPTF 140
DB KLSSYKPIKNTYNAKNTYPPVYKPKTYPTPKPSYPPTYK-----PKPSY 135
QY 141 KDVRSORFGFKYTKINKTPLPPLPPRAITLPKKYQPLPPAPPESSAYFAPKPTFFEVQR 200
DB 136 PATYKSKSYSPSSYKPKTYPTPKLTYPTTYPKPSYPPTYK-----YKPKPSYPATYK 190
QY 201 G-----PRORSKDFSRVLGAEEESHHTQTKPESSCPSSNONTOKSP-----AIASSSYM 250
DB 191 SKSSYPSPSYKTKTYTPSSYKPKTYPTPKVSYPTPKSKSYPIYTKASYPSYK 250
QY 251 PGKHSIQARDHTGSMQHCAPQRCQAASHSPRLPYENTNSEKPD--PT---KP 299
DB 251 PKK--TYPSTYKPKISYPPPTYKAKPSVPTSYRKPSPSPSYKAKPSYPPTYKAKP 303

Search completed: March 22, 2006, 01:36:41
Job time : 43 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2006, 01:28:54 ; Search time 233 Seconds
(without alignments)
1317.188 Million cell updates/sec

Title: US-10-717-619-2
Project score: 2316
Sequence: 1 MTSQGNKRTTKGFGDLRFQ.....QPLPLARLLLTQYSSQALHE 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2316	100.0	435	Q9JMJ3_MOUSE	Q9JMJ3 mus musculus
2	2305	99.5	435	Q9QZE2_MOUSE	Q9QZE2 mus musculus
3	2070	89.4	392	Q8C479_MOUSE	Q8C479 mus musculus
4	1306.5	56.4	428	Q7Z7G1_HUMAN	Q7Z7G1 homo sapien
5	1132	48.9	376	Q9P2U9_HUMAN	Q9P2U9 homo sapien
6	412.5	17.8	533	1 LCP2_MOUSE	Q60787 mus musculus
7	410.5	17.7	533	2 Q922M0_MOUSE	Q922M0 mus musculus
8	410	17.7	525	2 Q920L0_RAT	Q920L0 rattus norv
9	409	17.7	534	2 Q5SU7_MOUSE	Q5SU7 mus musculus
10	398	17.2	533	1 LCP2_HUMAN	Q13094 homo sapien
11	398	17.2	533	2 Q53XV4_HUMAN	Q53XV4 homo sapien
12	381	16.5	530	2 Q9DG07_CHICK	Q9DG07 gallus gall
13	360	15.5	501	2 Q5U3P8_BRARE	Q5U3P8 brachydanio
14	359	15.5	508	2 Q6GP30_XENLA	Q6GP30 xenopus lae
15	355	15.3	501	2 Q8J136_BRARE	Q8J136 brachydanio
16	331	14.3	523	2 Q8C8A1_MOUSE	Q8C8A1 mus musculus
17	329.5	14.2	524	2 Q5SUP8_MOUSE	Q5SUP8 mus musculus
18	319	13.8	456	2 Q75498_HUMAN	Q75498 homo sapien
19	319	13.8	456	2 Q8WV28_HUMAN	Q8WV28 homo sapien
20	315.5	13.6	552	2 Q9YGC1_CHICK	Q9YGC1 gallus gall
21	309.5	13.4	433	2 Q75499_HUMAN	Q75499 homo sapien
22	308	13.3	544	2 Q5HZD8_XENTR	Q5HZD8 xenopus tro
23	307.5	13.3	457	2 Q4KM52_RAT	Q4KM52 rattus norv
24	304.5	13.1	529	2 Q8JG66_XENLA	Q8JG66 xenopus lae
25	299	12.9	549	2 Q5U4V8_XENLA	Q5U4V8 xenopus lae
26	297.5	12.8	457	2 Q9QUN3_MOUSE	Q9QUN3 mus musculus
27	295.5	12.8	490	2 Q8J135_BRARE	Q8J135 brachydanio
28	287.5	12.4	457	2 Q88504_MOUSE	Q88504 mus musculus
29	219.5	9.5	297	2 Q9D413_MOUSE	Q9D413 mus musculus
30	194.5	8.4	175	2 Q7Z4S9_HUMAN	Q7Z4S9 homo sapien
31	188.5	8.1	516	2 Q54737_MOUSE	Q54737 mus musculus

32	165.5	7.1	594	1	SHC3_HUMAN	Q92529 homo sapien
33	165.5	7.1	594	2	Q5T7I7_HUMAN	Q5T7I7 homo sapien
34	151	6.5	467	2	Q68EJ7_BRARE	Q68EJ7 brachydanio
35	150	6.5	392	2	Q5RCZ3_PONPY	Q5RCZ3 pongo pygma
36	148	6.4	393	2	Q619C6_CAEBR	Q619C6 caenorhabdi
37	147.5	6.4	2511	2	Q5CRQ0_CRYPV	Q5CRQ0 cryptospori
38	146	6.3	724	1	P85B_HUMAN	P23726 bos taurus
39	146	6.3	728	1	P85B_HUMAN	O00459 homo sapien
40	146	6.3	728	2	Q5EAT5_HUMAN	Q5EAT5 homo sapien
41	145.5	6.3	722	2	Q5FVS6_RAT	Q5FVS6 rattus norv
42	143.5	6.2	568	2	Q5PBF8_GIBFU	Q5PBF8 gibberella
43	143.5	6.2	574	2	Q8IAA6_CAEL	Q8IAA6 caenorhabdi
44	143.5	6.2	662	2	Q9N3S5_CAEL	Q9N3S5 caenorhabdi
45	143.5	6.2	700	2	Q7YZX2_SUBDO	Q7YZX2 suberites d

ALIGNMENTS

RESULT 1
Q9JMJ3_MOUSE
ID Q9JMJ3_MOUSE PRELIMINARY; PRT; 435 AA.
AC Q9JMJ3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MIST.
GN Names=Clnk; Synonyms=MIST;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20288989; PubMed=10744659; DOI=10.1093/intimm/12.4.573;
RA Goiteinuk R.; Kanazashi H.; Sasanuma H.; Fujimura Y.; Hataka Y.;
RA Tatsuno A.; Ra C.; Hayashi K.; Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein MIST/Clnk involved in IgE
RT receptor-mediated mast cell degranulation.";
RL J. Int. Immunol. 12:573-580(2000).
DR EMBL; AB027320; EMBL56240.1; -, mRNA.
DR HSSP; P00524; INZL.
DR MGI; MGI:1351468; Clnk.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; IDA.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
SQ SEQUENCE 435 AA; 49513 MW; 4D1BD3E2F0C61ED6 CRC64;

Query Match	100.0%;	Score	2316;	DB 2;	Length	435;			
Best Local Similarity	100.0%;	Pred. No.	3.4e-134;						
Matches	435;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MTSQGNKRTTKGFGDLRFQNVSLLNKRSWPSLSAKGRCRAVLEPLPDHRRNLAVPGG	60						
Db	1	MTSQGNKRTTKGFGDLRFQNVSLLNKRSWPSLSAKGRCRAVLEPLPDHRRNLAVPGG	60						
Qy	61	EKCSNNNDYEDBFFQLLKAWPSMKIILPARPIQSEYADTRYFQDMWEAPLLLPKASVST	120						
Db	61	EKCSNNNDYEDBFFQLLKAWPSMKIILPARPIQSEYADTRYFQDMWEAPLLLPKASVST	120						
Qy	121	EROTRDVRMTQLBEVDKPTFKDVRQSRQFKGKYTKTKNTKPLPPRPAITLPPKYQPLPPA	180						
Db	121	EROTRDVRMTQLBEVDKPTFKDVRQSRQFKGKYTKTKNTKPLPPRPAITLPPKYQPLPPA	180						
Qy	181	PPESAYFAPKPTFFVQVQRPQSRQSAKQFSRVLGAEEESHHTQKPSSCPSSNQNTQKS	240						
Db	181	PPESAYFAPKPTFFVQVQRPQSRQSAKQFSRVLGAEEESHHTQKPSSCPSSNQNTQKS	240						

```

QY 241 PPAIASSYMPGKHSIQARDHTGSMQHCPCPAORCQAAASHSPRLPYENTNSEKPDPTKPD 300
Db 241 PPAIASSYMPGKHSIQARDHTGSMQHCPCPAORCQAAASHSPRLPYENTNSEKPDPTKPD 300
QY 301 EKDVQNEWYIGYSQAQVEDVLKMKDKGTFLVRDCSTKSKAEPYVLVVFYGNKVNVK 360
Db 301 EKDVQNEWYIGYSQAQVEDVLKMKDKGTFLVRDCSTKSKAEPYVLVVFYGNKVNVK 360
QY 361 IRFLESNQFALGTGLRGNEFMFSDVEDIIEHYTFYFILLIDGDKAARRKQCYLTQPLPL 420
Db 361 IRFLESNQFALGTGLRGNEFMFSDVEDIIEHYTFYFILLIDGDKAARRKQCYLTQPLPL 420
QY 421 ARLLLTQYSSQALHE 435
Db 421 ARLLLTQYSSQALHE 435

RESULT 2
Q9QZE2 MOUSE
ID Q9QZE2_MOUSE PRELIMINARY; PRT; 435 AA.
AC Q9QZE2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CLNK.
GN Names=Clnk;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=20029811; PubMed=10562326; DOI=10.1084/jem.190.10.1527;
RA Cao M.Y., Davidson D., Yu J., Latour S., Veilleux A.;
RT "Clnk, a novel SLP-76-related adaptor molecule expressed in cytokine-
stimulated hemopoietic cells."
RL J. Exp. Med. 190:1527-1534(1999).
DR EMBL; AF187819; AAF14299.1; -; mRNA.
DR HSSP; P00524; INZL.
DR Ensembl; ENSMUSG0000039315; Mus musculus.
DR MGI; MGI:1351468; Clnk.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IDA.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 435 AA; 49492 MW; 5CD27EC971FC0EA5 CRC64;

Query Match 99.5%; Score 2305; DB 2; Length 435;
Best Local Similarity 99.5%; Pred. No. 1.6e-133;
Matches 433; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTSQGNKRTTKGFGDLRFQNVSLLNKRSWPSLSAGKRCRAVLEPLPDHRRNLAVPGG 60
Db- 1 MTSQGNKRTTKGFGDLRFQNVSLLNKRSWPSLSAGKRCRAVLEPLPDHRRNLAVPGG 60
QY 61 EKCSNNYDEDPFQLLKAWPSMKILPARIQSEYADTRYFQDMWEAPLLLPKASVST 120
Db 61 EKCSNNYDEDPFQLLKAWPSMKILPARIQSEYADTRYFQDMWEAPLLLPKASVST 120
QY 121 ERQTRVMTQLEVDKPFKVDRSQRFKTKTKNKTPPLPPRAITLPKKYQPLPPA 180
Db 121 ERQTRVMTQLEVDKPFKVDRSQRFKTKTKNKTPPLPPRAITLPKKYQPLPPA 180
QY 181 PPESAYFAPKPTTFPEVQGRPRQRAKQPSRVLGAEESHHQTKPESSCPSSNQNTQKS 240
Db 181 PPESAYFAPKPTTFPEVQGRPRQRAKQPSRVLGAEESHHQTKPESSCPSSNQNTQKS 240

```

```

RESULT 3
Q8C479 MOUSE
ID Q8C479_MOUSE PRELIMINARY; PRT; 392 AA.
AC Q8C479;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus ES cells cDNA, RIKEN full-length enriched library, linker,
clone:C330034F12 product:cytokine-dependent hematopoietic cell linker,
full insert sequence. (Fragment).
GN Names=Clnk;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=99279283; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gibbs C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtauki S.,
Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).

```

```

[4]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=20493174; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,;
RA "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kajoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Nakazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK082826; BAC38640.1; -, mRNA.
DR HSP; P00524; INZL.
DR MGI; MGI:1351468; Clnk.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IDA.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
FT NON TER
SQ SEQUENCE 392 AA; 44749 MW; 29B5E68BECA75FD0 CRC64;

Query Match 89.4%; Score 2070; DB 2; Length 392;
Best Local Similarity 99.2%; Pred. No. 3.8e-119;
Matches 389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 44 LEPLDHRRLAGVPGGKCNNDYEDPEFQLLKAWPMSMKILPARIQSEYADTRYFQ 103
Db 1 LEPLDHRRLAGVPGGKCNNDYEDPEFQLLKAWPMSMKILPARIQSEYADTRYFQ 60

Qy 104 DMMEAPLLLPKASVSTRQTRDVRMTQLEVDKPTFKDVRQSRFGKFKYTKINKTPLPP 163
Db 61 DTMEAPLLLPKASVSTRQTRDVRMTQLEVDKPTFKDVRQSRFGKFKYTKINKTPLPP 120

Qy 164 PRPAITLPKKYQPLPPAPPESSAYFAPKPTFFPEVQGRQSRSAKDFSRVLGAEEESHQ 223
Db 121 PRPAITLPKKYQPLPPAPPESSAYFAPKPTFFPEVQGRQSRSAKDFSRVLGAEEESHQ 180

Qy 224 TKPSSCSPSSNONTQKSPPAIASSYMPGKHISIQARDHTGSMQCHPCAQRQAAASHSPRM 283
Db 181 TKPSSCSPSSNONTQKSPPAIASSYMPGKHISIQARDHTGSMQCHPCAQRQAAASHSPRM 240

Qy 284 LPYENTNSEKPDPTKPKDEKQVQNEWYIGYSRQAVEDVLMKENKDGTFVLVRDSTKSKA 343
Db 181 LPYENTNSEKPDPTKPKDEKQVQNEWYIGYSRQAVEDVLMKENKDGTFVLVRDSTKSKA 300
Db 344 EPVVLVVFYGNKYNKIRPLESNQOQFALCTGLRGNEFMFDSVEDIIIEHYTFYFPILLDGGK 403
Db 301 EPVVLVVFYGNKYNKIRPLESNQOQFALCTGLRGNEFMFDSVEDIIIEHYTFYFPILLDGGK 360
Qy 404 DKAARRKQCYLTQPLPLARLLLTQYSSQALHE 435
Db 361 DKAARRKQCYLTQPLPLARLLLTQYSSQALHE 392

RESULT 4
Qy 727761 HUMAN PRELIMINARY; PRT; 428 AA.
AC Q72761;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE MIST.
GN Name=mist;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Goitsuka R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB110420; BAC76765.1; -, mRNA.
DR HSP; Q64010; IJ05.
DR Ensembl; ENSG00000109684; Homo sapiens.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
SQ SEQUENCE 428 AA; 49482 MW; DC1015112AF460D1 CRC64;

Query Match 56.4%; Score 1306.5; DB 2; Length 428;
Best Local Similarity 61.7%; Pred. No. 3e-72;
Matches 264; Conservative 47; Mismatches 112; Indels 5; Gaps 4;

Qy 1 MTSQGNKRTTKGFGDLRFQNVSLLKORSWPSLSAAGRCRAVLPLPDHRRNLGAVPGG 60
Db 1 MNRQGNKRTTKGSGNDLKFNFLPKNRSWPRINSATGQYQRMNKPILLDWRNFVAALVDG 60

Qy 61 EKCNSNDYEDPEFQLLKAWPMSMKILPARIQSEYADTRYFQDMMEAPLLLPKASVST 120
Db 61 AKGHSDDDDYDPELRMEETWQSIKILPARIQSEYADTRYFQVAMDTPLDTRTSISI 120

Qy 121 ERQTRDVRMTQLEVDKPTFKDVRQSRFGKFKYTKINKTPLPPRAITLPKKYQPLPPA 180
Db 121 GQPTWNTQ-TRLEVRDVKPTISRVSQNKIGDASVRKNKIPLPPRPLITLPKKYQPLPP- 178

Qy 181 PPRESSAYFAPKPTFFPEVQGRQSRSAKDFSRVLGAEEESHQTKPSSCSPSSNONTQKS 240
Db 179 EPSSRPPLSRHTFFEVQMPQSIULRDLSEVLEAKVPHNQKPESTHLENQNTQEI 238

Qy 241 PPAIASSYMPGKHISIQARDHTGSMQCHPCAQRQAAASHSP--RMLPYENTNSEKPDPTK 298
Db 239 PLAISSTTSNHSVQNRDHRGDMQPCSPQRCPPASCSPHENILPKYKTSWRPFPKR 298

Qy 299 PDEKDVQNEWYIGYSRQAVEDVLMKENKDGTFVLVRDCTSKAEPVVLVVFYGNKYN 358
Db 299 SDRKDVQNEWYIGYSRQAVEAFMKNKDGSLVRDCTSKKEEYVLAIFYENKYN 358

Qy 359 VKIRPLESNQOQFALCTGLRGNEFMFDSVEDIIIEHYTFYFPILLDGGK-AARRKQCYLTQ 417
Db 359 VKIRPLESNQOQFALCTGLRGNEFMFDSVEDIIIEHYKNFPIILIDGKDTGVRKQCHLTQ 418
```

```
QY 418 LPLARLL 425
Db 419 LPLRHL 426

RESULT 5
O9P2U9 HUMAN PRELIMINARY; PRT; 376 AA.
AC Q9P2U9
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MIST (Fragment).
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20208989; PubMed=10744659; DOI=10.1093/intimm/12.4.573;
RA Goitsuka R., Kanazashi H., Sasanuma H., Fujimura Y., Hidaka Y.,
RA Tateono A., Ra C., Hayashi K., Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein MIST/Clnk involved in Ige
RT receptor-mediated mast cell degranulation.";
RL Int. Immunol. 12:573-580(2000).
DR EMBL; AB032369; BAA96241.1; -; mRNA.
DR HSSP; Q06124; 2SHP.
DR GO; GO:0005622; C:intracellular; NAS.
DR GO; GO:0005070; F:SH3/SH2 adaptor activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR GO; GO:0007242; P:intracellular signaling cascade; IDA.
DR InterPro; IPR000980; SH2.
DR PRINTS; PR00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
FT NON TER 1
FT NON TER 376
SQ SEQUENCE 376 AA; 43542 MW; 389421B629B028B2 CRC64;

Query Match 48.9%; Score 1132; DB 2; Length 376;
Best Local Similarity 60.9%; Pred. No. 1.3e-61;
Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

QY 19 FQNVSLKNSWPSLSAKGRCRAVLEPLPDHRLNLAGVPGGKCNNDYEDPEFQLLK 78
Db 1 FQNFSLPKNSWPRINSATGQYQRMKPLLDWERNFAAULDGAKGSHDDDDYDFELRMBE 60

QY 79 AWPMSKILPARPTQSEYADTRYFQDMWEAPLLLPKASVSTERQTRDVRMTQLEEVDPK 138
Db 61 TWQSIKILPARPIKESEYADTHYFKVAMDTPLDTRTISIGQPTWNTQ-TELERVDKP 119

QY 139 TFKDVRSQRFKPKYTKINTPLPPRPAILTPKKYQIPLPPAPPESSAYFAPKPTFPEV 198
Db 120 ISRDVRSQNIKGDAVSRKNKIPLPFRPLITLPKKYQLPLP-EPESRRPPLSQRHTFPEV 178

QY 199 QRGPRQSRSAKDFSRVLGAEESHQTKPSSSCSSNQTOKSPATASSYMPGKHSIOA 258
Db 179 QGFPQSISURDLSEVLAEKVPINQRPSTHLENNQNTQEIPLAISSSFTTSNHSVQN 238

QY 259 RDHTGSMQHCAPQRCQAAASHSP--RMLPYENTNSSEKPDPTKPEKDVQWQNEYIGEYSR 316
Db 239 RDHRGGMQPCSPQRCQPPASCSPHENILPKYTSWRPPPKRSDRKDVQWQNEYIGEYSR 298

QY 317 QAVEDVLMKENKDGTFVLVDCSTKSKAEPYVLVVFYGNKYNVKIRPLESNQOFPALGTGL 376
Db 299 QAVEEAFMKENKDGSLVLRDCSTKSKPEPYVLAVFYENKYNVKIRFLERNQOFPALGTGL 358

QY 377 RGNEMFDSVEDIIEHY 392
Query Match 17.8%; Score 412.5; DB 1; Length 533;
```


RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [6]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Bone;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Akizawa J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1617-1630 (2000).
 RN [7]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Bone;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Konno H., Akizawa J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [8]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Bone;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Konno H., Akizawa J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [8]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Bone;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Konno H., Akizawa J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771 (2000).
 RN [8]

Query Match 17.7%; Score 410.5; DB 2; Length 533;
 Best Local Similarity 29.7%; Pred. No. 3.9e-17;
 Matches 148; Conservative 50; Mismatches 176; Indels 125; Gaps 20;

QY 14 FGDLRFONVSL-----KNSWPSLSAKRCRAVLEPLPDHRRNLG-----56
 DB 57 FPKLRPLLSKLSQDINKNEERSISIFRKQIPRFLEETESHEDDGGSFFEDDYESP 116
 QY 57 --VPGKCNKNNYEDP--EFOLL-----KAWPS-----MKILPARPTQ---ESEY 96
 DB 117 DDDPDGE---DDGDYSPNEEQALVDDAADYPPPPSNNEALQSSILPPNSFHTNSMY 173
 QY 97 ADTRYFQDMMEAPLLLP--PKASV-----EESAYFAPKPTFPVQRPQRSKADF 210

DB 174 IDRPPTGKVSQQPPVPLRPKLPALPPTGRNHSPLSPHPNHEEBSRSNNKTAKLPAP 233
 QY 119 STEROTRDVRMTQLEEDV-----KPTFKDVSQRFGKFKYTKINKTLPPLPPPAITL 170
 DB 234 SIDRSTKPLDRLSLAPLDREPFLTKKPFSDKPSAPLGRHLPKIQKPLP---PAMDR 290
 QY 171 PKYQDPLPAPPRESSAYFAPKPTFFEVQVQRPQSAKDSRVLGABEESHQTKPSSC 230
 DB 291 HERNERLGPVTR-----KPPVPRHGRGDPORRE-----NDEDDVHQRLPQPSL 334
 QY 231 PSNQNTQKSPPAIASS-----SYMPGKHISIQARDHTGSMQHC--PAQRCQAAASHP- 281
 DB 335 PMSSTNTPFSRSVQPSKNTFFPLAHMPGAFS-----ESNIGFQQASALPPYFSQSPGNRPL 391
 QY 282 -----RMLPYENTNSEKPDPTKPDKDV-WQNEWYIGEYSQAVEDVLMKENKDGTFVLVRD 336
 DB 392 RSEGRNLPLVPV--RPQPSGEEETPLDEEYVSVITRPEAEALRKINQDGTFLVRD 449
 QY 337 CSTKSKABPYLVVFFGVNKNYKIRFLESNQFALGTGLRGNEMFDSVEDIEHYTFP 396
 DB 450 SSKKTANNPVLMVLYKDYNIQIRYQESQVYLLGTGLRKEDEFLSVSDIIDYFRMP 509
 QY 397 ILLIDGKDKAARRKQCVLT 415
 DB 510 LLLIDGNRGS-RYQCTLT 527

RESULT 8

Q920L0_RAT PRELIMINARY; PRT; 525 AA.
 ID Q920L0_RAT PRELIMINARY; PRT; 525 AA.
 AC Q920L0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE SLP-76 adaptor protein.
 GN Name=SLP2; Synonyms=SLP76;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Fujii Y., Goithuka G., Wakahara S., Hara T., Nakao T.;
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB072980; BAB71779.1; -; mRNA.
 DR HSSP; Q60787; IOEB.
 DR Ensembl; ENSRNOG00000005620; Rattus norvegicus.
 DR RGD; 619743; Lcp2.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR InterPro; IPR011510; SAM_2.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF07647; SAM_2; 1.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00454; SAM; 1.
 DR PROSITE; PS50001; SH2; 1.
 SQ SEQUENCE 525 AA; 59521 MW; 0B1CDD40DAE7372 CRC64;

Query Match 17.7%; Score 410; DB 2; Length 525;
 Best Local Similarity 31.9%; Pred. No. 4.1e-17;
 Matches 127; Conservative 51; Mismatches 148; Indels 72; Gaps 16;
 QY 68 DYEDPFFQLKAWPSMKILPARPI--QSEYADTRYFQDMMEAP-----LLLP--- 114
 DB 144 DYEPPPSNDEEALQN-SILPAKFPNTNSMYIDRPPTGKVSQQPPVPPQRPMAALPLPT 202
 QY 115 -----KASVSTRQTRDVRMTQLEEDVKTDFKDVRSQRFGKFKYTKINKTPL---- 161
 DB 203 GRNHSNHEEBSRSNRNHTAKLPAPSIDRSTKPLD--RSLA-----PPLDREPLLGKK 254
 QY 162 PPRPRAITLPKKYQP---LPPAPP-----EESAYFAPKPTFPVQRPQRSKADF 210

Db	255	: :
----	-----	---

Db	291	HERNERLGPVTR-----KPPVPRHGRGPDORRE-----NDEDDVHORPLPQPSL	334
Qy	231	PSNQNTQKSPPAIASS-----SYMPGKHSIQARDHTGSMQHC--PAQRCAAAASHSP-	281
Db	335	PSMSNSTFPSPSRVQSPSSKNTFPLAHMFGAFS---ESNIGFQQSASLPPYFSQEGPGNRPP	391
Qy	282	-----RMLPYENTNSEKPDPTKDEKDV-QWNEWYIGEYSRQAVEDVLAKENKDGTLVR	335
Db	392	LRSEGRNLPLVPVN--RPQPPSPGEEETPLDEWYVSYITRPEAEALARKINQDGTFLVR	449
Qy	336	DCSTKSKAEPIVLVVFYGNKVNKIRFLESNQOQFALGTLRGNEMFDSVEDIEHYTFP	395
Db	450	DSKKTANNPVLVWLYKDKVNIQIRYQESQVILLGTGLRGKEDFLSVSDIIDYFRKM	509
Qy	396	PILLIDGKDKAARRKQCYLT	415
Db	510	PLLLIDGKNRGS-RYQCTILT	528
RESULT 10			
LCP2 HUMAN			
ID	LCP2	HUMAN STANDARD;	PRT; 533 AA.
AC	Q13034;		
DT	16-OCT-2001	(Rel. 40, Created)	
DT	16-OCT-2001	(Rel. 40, Last sequence update)	
DT	10-MAY-2005	(Rel. 47, Last annotation update)	
DE	Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).		
GN	Name=LCP2;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.		
RC	TISSUE=Leukemia;		
RX	MEDLINE=95221345; PubMed=7706237; DOI=10.1074/jbc.270.13.7029;		
RA	Jackson J.K., Motto D.G., Sun Q., Tanemoto M., Turck C.W., Peltz G.A.,		
RA	Koretzky G.A., Findell P.R.;		
RT	"Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with Grb2 in T cells."		
RL	J. Biol. Chem. 270:7029-7032(1995).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].		
RC	TISSUE=Prostate;		
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,		
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
CC	-!- FUNCTION: Involved in T cell antigen receptor mediated signaling.		
CC	-!- SUBUNIT: Interacts with SLA (By similarity). Interacts with the adapter proteins GRB2 and Fyb.		
CC	-!- INTERACTION:		
CC	P08631:HCK; NExp=1; IntAct=EBI-346946, EBI-346340;		

CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -|- TISSUE SPECIFICITY: Highly expressed in spleen, thymus, and
 CC peripheral blood leukocytes. Highly expressed also in T-cell and
 CC monocytic cell lines, expressed at lower level in B-cell lines.
 CC Not detected in fibroblast or neuroblastoma cell lines.
 CC -|- PTM: Phosphorylated after T-cell receptor activation by ZAP-70.
 CC -|- SIMILARITY: Contains 1 SH2 domain.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; U20158; AAC50135.1; -; mRNA.
 CC EMBL; BC016618; AAH16618.1; -; mRNA.
 CC PIR; A56110.
 CC PDB; 1H3H; NMR; B=232-241.
 CC IntAct; Q13094; -.
 CC Ensembl; ENSG00000043462; Homo sapiens.
 CC HGNC; HGNC:8529; LCP2.
 CC H-InvDB; HIX0005403; -.
 CC MIM; 601603; -.
 CC GO; GO:0005515; F:protein binding; TAS.
 CC GO; GO:0006955; P:immune response; TAS.
 CC GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; TAS.
 CC InterPro; IPR001660; SAM.
 CC InterPro; IPR011510; SAM_2.
 CC InterPro; IPR000980; SH2.
 CC Pfam; PF07647; SAM_2; 1.
 CC Pfam; PF00017; SH2_1.
 CC PRINTS; PR04041; SH2DOMAIN.
 CC ProDom; PD000093; SH2; 1.
 CC SMART; SM00454; SAM; 1.
 CC SMART; SM00252; SH2; 1.
 CC PROSITE; PS50001; SH2; 1.
 CC 3D-structure; Direct protein sequencing; Phosphorylation; SH2 domain.
 FT DOMAIN 422 530
 FT COMPBIAS 133 136 Poly-Glu.
 FT COMPBIAS 198 201 Poly-Pro.
 SQ SEQUENCE 533 AA; 60188 MW; C5D22F31D36200C8 CRC64;
 Query Match 17.2%; Score 398; DB 1; Length 533;
 Best Local Similarity 30.2%; Pred. No. 2.3e-16;
 Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;
 QY 68 DYEDPEFQLLKAWPSMKILPARIQSE--YADTRYFQDMMEAP-----LLLPPKA 116
 DB 144 DYEPPPSNDEEALQN-SILPAKFPFNSNSMYIDRPPSGKTPQPPVPQRPMAALPPPPA 202
 QY 117 -----SVSTEROTRDVRMTQLEVD-----KPT 139
 DB 203 GRNHSPLPPQTNEEPPSRNHNKAKLPAPSIDRSTKPLDRSLAPFPDREPFTLGKKPP 262
 QY 140 FKD---VRSORFGKYTKINKTLPPLPPPAITLPKKYQPLPAPPESSAYFAPKPTFP 196
 DB 263 FSDKPSIPAGRSIGELHPKIQKPLPP-----TTERHERSSSLP-----GKKPPVP 308
 QY 197 EVQGRQRQSAKDFSVLGAEESESHQTKPE-----SSCPSSNQNTQKSPPAIASSSYM 250
 DB 309 KHGWGPDRE-----NDEDDVHORPLPQALLPMSSNTFFPSRSTKSPMNPPLSSH 360
 QY 251 PGKHSIQARDHTGSMQHCQAQRCQAASHSP-----RMLPYENTNKGKDPPTKPDKDVW 305
 DB 361 PGAFS-ESNSFPQSASLPYFSGPSNRPPRAEGRNFLPLPNKPRP-PSPAEENS 418
 QY 306 QNEWYIGEYSRQAVEDVLMKENKDGTFVLRDCSTKSKAEPLYLVVFGNKNVYNKIRFLE 365
 DB 419 NEEWYYSYITRPEAEALRKINQDGFVLRDSSKTTTNPYVLMVLVKDKVYNIQIRYQK 478
 QY 366 SNOQFALGTGLRGNEFDSVEDIIHYTFYFILLIDGKDKAARRKQCYLT 415
 DB 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLLLIDGKNRGS-RYQCTLT 527
 RESULT 12
 Q9DG07_CHICK PRELIMINARY; PRT; 530 AA.
 ID Q9DG07_CHICK
 AC Q9DG07;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE SLP-76 adaptor protein.
 GN Name=SLP76;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.

RESULT 11
 Q53XV4_HUMAN PRELIMINARY; PRT; 533 AA.
 ID Q53XV4_HUMAN
 AC Q53XV4;
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
 DE Lymphocyte cytosolic protein 2 (SH2 domain containing leukocyte
 DE protein of 76kDa).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
 RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
 RA Phelan M., Farmer A.;
 RT "Cloning of human full-length CDSs in BD Creator(TM) System Donor
 RT vector";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BT007273; AAP35937.1; -; mRNA.
 SQ SEQUENCE 533 AA; 60188 MW; C5D22F31D36200C8 CRC64;
 Query Match 17.2%; Score 398; DB 2; Length 533;
 Best Local Similarity 30.2%; Pred. No. 2.3e-16;
 Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;
 QY 68 DYEDPEFQLLKAWPSMKILPARIQSE--YADTRYFQDMMEAP-----LLLPPKA 116
 DB 144 DYEPPPSNDEEALQN-SILPAKFPFNSNSMYIDRPPSGKTPQPPVPQRPMAALPPPPA 202
 QY 117 -----SVSTEROTRDVRMTQLEVD-----KPT 139
 DB 203 GRNHSPLPPQTNEEPPSRNHNKAKLPAPSIDRSTKPLDRSLAPFPDREPFTLGKKPP 262
 QY 140 FKD---VRSORFGKYTKINKTLPPLPPPAITLPKKYQPLPAPPESSAYFAPKPTFP 196
 DB 263 FSDKPSIPAGRSIGELHPKIQKPLPP-----TTERHERSSSLP-----GKKPPVP 308
 QY 197 EVQGRQRQSAKDFSVLGAEESESHQTKPE-----SSCPSSNQNTQKSPPAIASSSYM 250
 DB 309 KHGWGPDRE-----NDEDDVHORPLPQALLPMSSNTFFPSRSTKSPMNPPLSSH 360
 QY 251 PGKHSIQARDHTGSMQHCQAQRCQAASHSP-----RMLPYENTNKGKDPPTKPDKDVW 305
 DB 361 PGAFS-ESNSFPQSASLPYFSGPSNRPPRAEGRNFLPLPNKPRP-PSPAEENS 418
 QY 306 QNEWYIGEYSRQAVEDVLMKENKDGTFVLRDCSTKSKAEPLYLVVFGNKNVYNKIRFLE 365
 DB 419 NEEWYYSYITRPEAEALRKINQDGFVLRDSSKTTTNPYVLMVLVKDKVYNIQIRYQK 478
 QY 366 SNOQFALGTGLRGNEFDSVEDIIHYTFYFILLIDGKDKAARRKQCYLT 415
 DB 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLLLIDGKNRGS-RYQCTLT 527
 RESULT 12
 Q9DG07_CHICK PRELIMINARY; PRT; 530 AA.
 ID Q9DG07_CHICK
 AC Q9DG07;
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE SLP-76 adaptor protein.
 GN Name=SLP76;
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.

```

OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Thymus;
RX MEDLINE=20451093; PubMed=10993915; DOI=10.1084/jem.192.6.847;
RA Iehiai M., Kuroaki M., Inabe K., Chan A.C., Sugamura K., Kuroaki T.;
RT "Involvement of LAT, Gads, and Grb2 in compartmentation of SLP-76 to
RL the plasma membrane.";
DR J. Exp. Med. 192:847-856(2000).
DR EMBL: AF226988; AAG18493.1; -; mRNA.
DR HSSP: Q60787; 10EB.
DR Ensembl: ENSGALG00000002113; Gallus gallus.
DR GO: 0007242; P:intracellular signaling cascade; IEA.
DR InterPro: IPR001660; SAM.
DR InterPro: IPR011510; SAM 2.
DR Pfam: PF07647; SAM 2; 1.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00454; SAM; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 530 AA; 59580 MW; CF6CBCEBD79909A CRC64;

Query Match 16.5%; Score 381; DB 2; Length 530;
Best Local Similarity 29.8%; Pred. No. 2.5e-15;
Matches 123; Conservative 54; Mismatches 151; Indels 85; Gaps 16;

Qy 61 EKCSNND-YEDPEFQLLKAWPSMKILPARI-OESEVADTRYQDMMEAPLLLPKASV 118
Db 136 EEAHSDGVEYPPSPNNDEAHNV-IPPAKSLANNTYIDRPPTSRSHPQVPPQPGP 194
Qy 119 STERQTRDVRMTQLEEVDPKTFVDVRSQRFKGYTKINKTLPPLPPPAITL----- 170
Db 195 SPAPASFGRCASL-----PAFPLPGNNDENVK-----PSKPPAPSIDRSTKPLDR 242
Qy 171 -----PKK-----YQPLPPAPPESSAYPAPKPTPE- 197
Db 243 LGPPFERENPVGRKPGHPEKLLTFLQALGEQLAMWPKPVPVPSDRYGRNPSPLRKQI 302
Qy 198 -VQGGPQRSNAKSRVLGAEEESH--HQTKPESSCPSSNQNT-----QKSPPAIASSSY 249
Db 303 PVKQGAQQRKPE-----EEEDHIFQRAVQIISLPYSSNTSPFSKIAKPPKGSNS- 354
Qy 250 MPKGHSIQARDHTGSMQHCPCAPQRCQAAAASHPRM-----LPVENTNSEKDPPTKPE 301
Db 355 IPGASARSLSASGL---PPRFLGNNSRSPSGTADLRPLIPRQTAHQNTTEDE 411
Qy 302 KDVMQNEWYIGEYGRQAVEDVLMKENKDGTFVLVDCSTKSKAEPIVVLVFGYGNKQVNVKI 361
Db 412 -DSLNDWEYVAYISKPEAEALRKINKKDGTFVLVDCSTKSKAEPIVVLVFGYGNKQVNIQI 470
Qy 362 RFLESNQFALGTGLRGNEFDSVEDIEHYTFPILLIDGKDKAARKQCYL 414
Db 471 RYQEQDQTYLLGTGLKGKEDFSSVAHIIDYFQRTPELLIDGKGRGS-RNQCVL 522

RESULT 13
QS03P8 BRARE PRELIMINARY; PRT; 501 AA.
AC QS03P8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Stp76.
GN ORFNames=zgc:101809;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=9555;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh P.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RN and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Olfactory epithelium;
RA Director MGC Project;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC085444; AAH85444.1; -; mRNA.
DR ZFIN: ZDB-GENE-041114-97; zgc:101809.
DR GO: 0007242; P:intracellular signaling cascade; IEA.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS00001; SH2; 1.
SQ SEQUENCE 501 AA; 55995 MW; SDB1620B9DSD57A90 CRC64;

Query Match 15.5%; Score 360; DB 2; Length 501;
Best Local Similarity 28.3%; Pred. No. 4.5e-14;
Matches 115; Conservative 57; Mismatches 141; Indels 94; Gaps 17;

Qy 61 EKCSNNDYEDPEFQLLKAWPS---MKILPARIQSEYADTRYQDMMEAPLLLPKAS 117
Db 136 EGSDSDNSYEPPESE-----PSEDKAQICPAKPMDCYIDNNRTRVNVRSQPPVPE- 188
Qy 118 VSTERQTRDVRMTQLEEVDPKTF---KDVRSQRFKGYTKINKTLPPLPPPAITLPKY 174
Db 189 -----RPGGPGSPVPVDRPSVGMNMRREER-----PQTPKRPAPAVDRSKPP 232
Qy 175 QPL-----PPA-----PPESSA-----YFAPKPTPEVQRG 201
Db 233 GTLDRIHPVAVAGRGTSLLDRINPFRSVNAAPFRCPAMEAPGDPMRIKPSLPP-SLG 291
Qy 202 PRORSKDF-----SRVLGAEEESHQTKPESS--CPSSNQNTQKSPPAIASSSYMPGK 253
Db 292 VR-RSASVTPGYSQNRHLQDRHEFDHTARQTSNTFPLHARNPSRPPGTHGQSFQ--- 347
Qy 254 HSTQARDHTGSMQHCPCAPQRCQAAAASH--SPRMLPYENTNSEKDPPTKPEKDVNQNEWYI 311
Db 348 -----TDSVNPSPSLPAKQEAQMTDHRRSARASGPPQDMGGRQD-----MDPAWYL 393
Qy 312 GEYSRQAVEDVLMKENKDGTFVLVDCSTKSKAEPIVVLVFGYGNKQVNVKIFLESNQFA 371
Db 394 GQITRGEAESCLRRVNDGTFLVDRSSNRSSNQPYTLVLVYQDKVYNIQR--RNQDRFM 451
Qy 372 LGTGLRGNEFDSVEDIEHYTFPILLIDGKDKAARKQCYLTQP 417
Db 452 LGTGMKASESFERSVDIVNQHKHTPELLIDAKNRESGQKQCALIYP 498

RESULT 14

```

Query Match	15.5%;	Score 359;	DB 2;	Length 508;
Best Local Similarity	29.1%;	Pred. No. 5.3e-14;		
Matches 116;	Conservative 56;	Mismatches 145;	Indels 82;	Gaps 17;
Qy	64	NSNNDYEDPEFQLLKAWPSMKILPAP--IQESEYADTYFYQDMMEAPLLLPKA-----	116	
Db	137	DGNGDYEPDPPIEDTHPA--ICQIKPATNSPYADREFGSMKNPP--EPKPKPGNDAP	192	
Qy	117	-----SVSYTERQTRVRM--TQLEEVDDKPTF-----KQVRSQRPRGFKYTK	155	

```
Qy 254 HSIQARDHTGSMQHCPAQRCAAAASH--SPRMLPYENTNSEKPDPTKPDDEKDVWQNEWYI 311
Db 348 -----TDSVNFPSRSLPAKLQEAHTDHRRSARASPPQTDMGGRQD-----MDPAWYL 393
Qy 312 GEYSRQAVEDVLMKENKDGTFVLVDCSTKSKAEPYVLVVFYGNKVYVVKIRFLESNOQFA 371
Db 394 QOITRGEAESCLRRVNRDGFVLVRDSSNRFSNQPYTLVLVLYQDKVYNIQIR--RNQDRFM 451
Qy 372 LGTGLRGNEPFDSVEDIIEHYTFPILLIDGKOK-AAARRKOCYLTOP 417
Db 452 LGTGMKASESFERYSDIVNQHKHTPLLLIDAKNRESGQOKOCALIYP 498
```

Search completed: March 22, 2006, 01:35:55
Job time : 236 secs

This Page Deleted (2/2/10)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Rup on: March 22, 2006, 01:36:09 ; Search time 46 Seconds
(without alignments)
781.824 Million cell updates/sec

Title: US-10-717-619-2
Perfect score: 2316
Sequence: 1 MTSQGNKRTTKEGFDLRFQ.....QPLPARLLLTQYSSQALHE 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	435	2	US-09-856-061-2
2	1316.5	56.8	428	2	US-10-104-047-3344
3	1132	48.9	376	2	US-09-856-061-4
4	398	17.2	533	2	US-09-949-016-6629
5	391	16.9	602	2	US-09-949-016-10221
6	319	13.8	449	2	US-09-355-214-3
7	319	13.8	456	1	US-08-819-013-1
8	319	13.8	456	2	US-09-355-214-1
9	319	13.8	470	2	US-09-949-016-7976
10	287.5	12.4	457	2	US-09-355-214-5
11	213	9.2	84	2	US-09-513-999C-7233
12	165.5	7.1	474	2	US-08-729-416C-1
13	165.5	7.1	474	2	US-09-433-353-1
14	165.5	7.1	594	2	US-08-729-416C-7
15	165.5	7.1	594	2	US-09-433-353-7
16	146	6.3	529	2	US-09-949-016-8918
17	141.5	6.1	474	2	US-08-807-342B-2
18	140	6.0	469	1	US-08-353-550-6
19	140	6.0	469	1	US-08-551-687-6
20	137.5	5.9	474	2	US-08-729-416C-11
21	137.5	5.9	474	2	US-09-433-353-11
22	137.5	5.9	594	2	US-08-729-416C-17
23	137.5	5.9	594	2	US-09-433-353-17
24	137	5.9	143	2	US-10-104-047-3542
25	136.5	5.9	605	2	US-09-949-016-8269
26	136.5	5.9	630	2	US-09-949-016-7270
27	135	5.8	272	2	US-09-270-767-45259

ALIGNMENTS

RESULT 1

US-09-856-061-2 2 US-09-949-016-11135 670 2 US-09-949-016-11135 Sequence 11135, A
; Sequence 2, Application US/09856061 675 2 US-08-426-509A-4 675 2 US-08-426-509A-4 Sequence 4, Appli
; Patent No. 6831151 675 2 US-08-232-545-4 675 2 US-08-232-545-4 Sequence 4, Appli
; GENERAL INFORMATION: 675 2 US-09-977-261-4 675 2 US-09-977-261-4 Sequence 4, Appli
; APPLICANT: GOITSUKA, RYO 675 4 PCT-US95-05008-4 675 4 PCT-US95-05008-4 Sequence 4, Appli
; TITLE OF INVENTION: A mast cell-specific adapter molecules and CDNAs thereof 675 2 US-08-807-342B-4 675 2 US-08-807-342B-4 Sequence 10423, A
; FILE REFERENCE: 2001-0614A/WMC/00653 675 2 US-09-949-016-10423 675 2 US-09-949-016-10423 Sequence 9020, Ap
; CURRENT APPLICATION NUMBER: US/09/856,061 675 2 US-09-949-016-6393 675 2 US-09-949-016-6393 Sequence 6393, Ap
; CURRENT FILING DATE: 2001-07-10 675 2 US-09-248-796A-15743 675 2 US-09-248-796A-15743 Sequence 15743, A
; PRIOR APPLICATION NUMBER: JP11-263778 675 2 US-08-807-342B-5 675 2 US-08-807-342B-5 Sequence 5, Appli
; PRIOR FILING DATE: 1999-09-17 675 2 US-08-167-035-2 675 2 US-08-167-035-2 Sequence 2, Appli
; NUMBER OF SEQ ID NOS: 4 675 2 US-08-208-887A-2 675 2 US-08-208-887A-2 Sequence 2, Appli
; SOFTWARE: Patentin Ver. 2.0 675 2 US-08-539-005-2 675 2 US-08-539-005-2 Sequence 2, Appli
; SEQ ID NO: 2 675 2 US-09-280-598-5 675 2 US-09-280-598-5 Sequence 5, Appli
; LENGTH: 435 675 2 US-09-417-197-49 675 2 US-09-417-197-49 Sequence 49, Appli
; TYPE: PRT 675 2 US-09-417-197-67 675 2 US-09-417-197-67 Sequence 67, Appli
; ORGANISM: Mus musculus
US-09-856-061-2

Query Match	100.0%	Score	2316	DB	2	Length	435
Best Local Similarity	100.0%	Pred. No.	5.8e-204				
Matches	435	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MTSQGNKRTTKEGFDLRFQNVSLLNKRSWPSLSAAGRCRAVLPLPDHRRNLGAVPGG	60				
Db	1	MTSQGNKRTTKEGFDLRFQNVSLLNKRSWPSLSAAGRCRAVLPLPDHRRNLGAVPGG	60				
Qy	61	EKCNNSNDYEDPFFQLLKAWPSMKILPAPRIQSEYADTRYFQDMMEAPLLPPKASVST	120				
Db	61	EKCNNSNDYEDPFFQLLKAWPSMKILPAPRIQSEYADTRYFQDMMEAPLLPPKASVST	120				
Qy	121	EROTRDRVMTQLEBEVDKPTFKVRSORFGKFKYTKINKTLPDPPRAITLTKYQPLPPA	180				
Db	121	EROTRDRVMTQLEBEVDKPTFKVRSORFGKFKYTKINKTLPDPPRAITLTKYQPLPPA	180				
Qy	181	PPRESSAYFAPKPTFFEVQVGRQPSAKDFSRVLGAEESHHOTKPESSCPSSNQNTQKS	240				
Db	181	PPRESSAYFAPKPTFFEVQVGRQPSAKDFSRVLGAEESHHOTKPESSCPSSNQNTQKS	240				
Qy	241	PPAIASSSYMPGKHSTQARDHTGSMQHCPCQRCQAASHSPRMLPYENTNSEKPDPTKPD	300				
Db	241	PPAIASSSYMPGKHSTQARDHTGSMQHCPCQRCQAASHSPRMLPYENTNSEKPDPTKPD	300				
Qy	301	EKDVQWQNEWYIGYSRQAVEDVLMKENKDGTFELVRDCSTKSKAEPPYLVVYFGKNVNVK	360				
Db	301	EKDVQWQNEWYIGYSRQAVEDVLMKENKDGTFELVRDCSTKSKAEPPYLVVYFGKNVNVK	360				
Qy	361	IRFLESNOQFALGTGLRGNEFMFSDVETIIHHYTFPILLIDGDKAARRKOCYLTQPLPL	420				


```
Db 361 IRFLESNQFALGTGLRGNEFMDSVEDIIHYTFFPIILLIDGDKDAARRKQCYLTQPLPL 420
Qy 421 ARLLITQYSSQALHE 435
Db 421 ARLLLTQYSSQALHE 435
RESULT 2
US-10-104-047-3344
; Sequence 3344, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3344
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3344
Query Match 56.8%; Score 1316.5; DB 2; Length 428;
Best Local Similarity 62.1%; Pred. No. 2.8e-112;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;
Qy 1 MTSQGNKRTTKEGFGDLRFQNVSLLNKRSWPSLSAAGRCRAVLLEPLPDHRRNLGAVPGG 60
Db 1 MNRQGNKRTTKEGSDNLDKFQNFSLPKNRSWPRINSATGQYQRMNKLPLDWRNFAAALDG 60
Qy 61 EKCNNSNDYEDPEFOLLKAWPSMKILPARPIQSEYADTRYFODMMEAPLPPKASVST 120
Db 61 AKGHSDDDDYDDPELRMEETWQSIKILPARPIKESEYADTHYFKVAMDTPPLDTRTSISI 120
Qy 121 ERQTRDVRMTQLEVDKPTFKDVRQRQPSAKDFSRVLGABEESHQTKPSSCPSSNQTKS 240
Db 121 GQPTWNTQ-TRLERVDKPIKSDVRSONIKGDASVRKNKILPLPPRPPLITLPPKYQPLPP- 178
Qy 181 PPESSAYFAPKPTFFPVQGRQPSAKDFSRVLGABEESHQTKPSSCPSSNQTKS 240
Db 179 EPSSRRPPLSORHTFFPEVQRMPSQISLRDLSEVLEAEKVPHNQRPSTHLENQNTQEI 238
Qy 241 PPAIASSYMPGKHSIQARQHTGSMOHCPCAPQCOAAASHSP--RMLPYENTNSEKPDPTK 298
Db 239 PLAISSSFTTSHSVQNRDRHGMQPCSPQRCQPAPSCSPHENILPYKTSWRPFPKR 298
Qy 299 PDEKVMQNEWYIGEYSRQAVEDVLMKENKDGTFVLVDCSTKSKAEPYVLVVFYGNKVN 358
Db 299 SDRKQVQHNWYIGEYSRQAVEAEAFMKNKDGSLVDCSTKSKKEEPYVLAVFYENKVN 358
Qy 359 VKIRFLESNQFALGTGLRGNEFMDSVEDIIHYTFFPIILLIDGDK-AARRKQCYLTQ 417
Db 359 VKIRFLESNQFALGTGLRGDEKFDSDVEDIIHYKKNFPIILLIDGDKQTVHRRKQCHLT 418
Qy 418 LPLARLL 425
Db 419 LPLTRHLL 426
RESULT 3
US-09-856-061-4
; Sequence 4, Application US/09856061
; Patent No. 6831151
; GENERAL INFORMATION:
; APPLICANT: GOITSUKA, Ryo
; TITLE OF INVENTION: A mast cell-specific adapter molecules and cDNAs thereof
; FILE REFERENCE: 2001-0614A/WMC/00553
; CURRENT APPLICATION NUMBER: US/09/856,061
```

```
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: JP11-263778
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-856-061-4
Query Match 48.9%; Score 1132; DB 2; Length 376;
Best Local Similarity 60.9%; Pred. No. 2e-95;
Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;
Qy 19 FQNVSLLNKRSWPSLSAAGRCRAVLLEPLPDHRRNLGAVPGGKCNNSNDYEDPEFOLLK 78
Db 1 FQNFSLPKNRSWPRINSATGQYQRMNKLPLDWRNFAAALDGAKGHSDDDDYDDPELRME 60
Qy 79 AWPMSKILPARPIQSEYADTRYFODMMEAPLPPKASVSTQTRDVRMTQLEVDK 138
Db 61 TWQSIKILPARPIKESEYADTHYFKVAMDTPPLDTRTSISIQPTWNTQ-TRLERVDK 119
Qy 139 TFKDVRSQREFKGYTKINKTLPPLPPPAITLPPKYQPLPPAPPESSAYFAKPTPEV 198
Db 120 ISRDVRSQNIKGDASVRKNKILPLPPRPPLITLPPKYQPLPP-EPSSRRPPLSORHTF 178
Qy 199 QRQPRORSKDFSRVLGABEESHQTKPSSCPSSNQTKSPPAIASSYMPGKHSIQ 258
Db 179 QCMPSQISLRDLSEVLEAEKVPHNQRPSTHLENQNTQEIPLAISSSFTTSHSVQ 238
Qy 259 RDHTGSMOHCPCAPQCOAAASHSP--RMLPYENTNSEKPDPTKDEKVMQNEWYIGEYS 316
Db 239 RDHRGGMQPCSPQRCQPAPSCSPHENILPYKTSWRPFPKRSDRKQVQHNWYIGEYS 298
Qy 317 QAVEDVLMKENKDGTFVLVDCSTKSKAEPYVLVVFYGNKVNKIRFLESNQFALGTGL 376
Db 299 QAVEEAFMKNKDGSLVDCSTKSKKEEPYVLAVFYENKVNKIRFLERNQFALGTGL 358
Qy 377 RGNEMFDSVEDIIHY 392
Db 359 RGDEKFDSDVEDIIHY 374
RESULT 4
US-09-949-016-6629
; Sequence 6629, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6629
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6629
Query Match 17.2%; Score 398; DB 2; Length 533;
Best Local Similarity 30.2%; Pred. No. 7.2e-28;
Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;
```


181	Qy	PREESSAYPAPKPTTPEVQGRGRORSKADPRVLGABEESSHOTKBPSSCPSSNQNT----	237
176	Db	EDNENYIHTPTSESSPPPEKAPM-----VNRSTKPNSSPTFASPGTAGS 219	
238	Qy	-----OKSPDPAIASSYMP--GKH-----SIQARDHTGSM--QHCPAORCOAA-- 276	
220	Db	RNSGAWETKSPPP--NAPSPPLPRAGKPTTPLKTTTPVASQNASVCEEKPIPAERRHSGS 278	
277	Qy	----ASHSPRMLPYENTNSKDPD-----TKPDEKDVQWQNE 308	
279	Db	HRQEAQVSVPFPFAQKHQKPIPLPRFTEGNGPTVDGPLPFSFNSTTISEQEAQGLCKP 338	
309	Qy	WTYIGESROAQVDVLKMKNGDGTFLVRDCSTKSAEYPVLVVFVGKNVYVVKIRFLESNO 368	
339	Db	WTAGACDRKSABEALHRKNKQGSFLIRKSSGHDSDKQPYTLVVFVFNKRVYVNPVRFIEATK 398	
369	Qy	QFALGTGLRGNEFMFDSVEDIIIEHYTYPFILLIDGK 404	
399	Db	QVALARKKNGGEYFGVSAEIRNHOHGPLVLIDSON 434	

RESULT 7

```

US-08-819-013-1
; Sequence 1, Application US/08819013
; Patent No. 5994522
; GENERAL INFORMATION:
; APPLICANT: Chan, Andrew C.
; TITLE OF INVENTION: BLNK PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,013
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/788,322
; FILING DATE: 24-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-64383-1/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-819-013-1

```

121	Qy	EROTRQVRMTQLEEVDPKPTFKDVRQSRQFKGYTKINKTPIPLPPRAIPATLPPKKYQPLPPA	180
145	Db	EKARLTSTLPALTALQKQV-----PPKPKGLLEDEADYVVPV	182
181	Qy	PPBESSAYFAPKPTTFPEVQRCPRORSAKDFRVLGABEESHQTKPESSCPSSNQNT----	237
183	Db	EDNDENYIHPTESSPPPEKAPM-----VNRSTKPNGSTPASPPTGTSAG	226
238	Qy	-----QKSPPAATASSYMP--CKH-----SIQARDHTGSM---OHCFAQRCCOA--	276
227	Db	RNSGAWTKSGPPP--AAPSPUPRAGKPTTPLKTTFVASQQNASSVCEBKPIPAERHRGSS	285
277	Qy	-----ASHSRMLPYENTNSEKPD-----TKPDKDVQWQNE	308
286	Db	HRQEAUVSPVFPFAQKQIHKPIPLPRFTEGNTVTDGPLPIFSSNSTIISEQAGVLCKP	345
309	Qy	WYIGEYGRQAVEDVLMKENKDGTFLVRDCSTKSKAEYPVLVVFYGNKRVYNNKIRLESNQ	368
346	Db	WYAGACDRKSAEBEALHRNKGDSFLIRKSSGHDQKQPYTLVVFFNKRVYNNIPVRIEATK	405
369	Qy	QFALGTGLRNMFPDSEDIIETHYTPILLIDKD	404
406	Db	QYALGRKNGEYFEGSVAEIIRNHQHSPLVIDSQN	441

RESULT 8

```

US-09-355-214-1
; Sequence 1, Application US/09355214
; Patent No. 6638724
; GENERAL INFORMATION:
; APPLICANT: Washington University
; TITLE OF INVENTION: BLNK PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/355,214
; FILING DATE: 23-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/819,013
; FILING DATE: 17-MAR-1997
; APPLICATION NUMBER: US 08/788,322
; FILING DATE: 24-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: FP-64383-2/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-355-214-1
Query Match 13.8% Score 319; DB 2; Length 456;

```

```

Qy 238 -----QKSPRAIASSYMP--GKH-----SIOARDHTGSM---OHCPAQRQAAA- 276
Db 241 RNSGAWETKSPPP-APSPPLPRAGKKTPTPLKTPPVASQQNASVCEEKIPPAERHKGSS 299
Qy 277 ----ASHSPRLMPEYNTNSEKDDP-----TKPDEKDVWQNE 308
Db 300 HRQEAQVSPVFPAPQKQIHQKPIPLPRFTBGGNPTVDGPLSPSSNSTISEQAGVLCCKP 359
Qy 309 WYIGEYSRQAVEDVLKMKENKDGTFILVRDCTSKSABEYVLWYVGNKVYVKNKIRFLESNQ 368
Db 360 WYAGACDRKSAAEALHRSNKGSLIRKSSGHDCKOPYTLLWPFNKRVYINPVRFIEATK 419
Qy 369 OFALGTCLGRNMFEDSVEDLIEHYTFYFPILLIDGKD 404
Db 420 QYALGRKKGGEYFGSVAEIIRNHQHSPLVLIDSQN 455

RESULT 10
US-09-355-214-5
; Sequence 5, Application US/09355214
; Patent No. 6638724
; GENERAL INFORMATION:
; APPLICANT: Washington University
; TITLE OF INVENTION: BLANK PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/355,214
; FILING DATE: 23-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/819,013
; FILING DATE: 17-MAR-1997
; APPLICATION NUMBER: US 08/788,322
; FILING DATE: 24-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: FP-64383-2/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 457 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-355-214-5

Query Match 12.4%; Score 287.5; DB 2; Length 457;
Best Local Similarity 25.5%; Pred. No. 7.9e-18;
Matches 106; Conservative 53; Mismatches 139; Indels 117; Gaps 16;

Qy 66 NNDYEPEFOLLKAWSPMKILPARPIQESYADTRYQDMMEAPLLLPKPAVSSTERQTR 125
Db 69 DSDYENPDEH---SDSEMYNPAETGDDSYE-----PPPA-----EQOTR 106

126 DVA-----MTOLBEVDKFTFKDVSASORFK-GFKYTKINKT--PLPPPPPAITLP-----KK 173

```

```
Db 107 VHPALPFTTGEYVDNR-----SQHSPPFSKTLPSKPSWPSAKARLASTLPAPNSLQK 161
Qy 174 YQLPPAPP---BESSAYFAPKPTFPVQVGRQRS-----AKDFSRVLGAEESHOT 224
Db 162 PQ-VPPKPKDLEADYVVDNDENYTHPRESSPPPAEKAPMWNRSKPNSSSKHMS 220
Qy 225 KPES-----SCSSNQNTOKSP-----PAIASSSYMPGKHSIQAR 259
Db 221 PPGTVAGRNQVWDSKSLPAAPSPUPRAGKKPATPLKTTVPVPLPNASNVCEKVPFAE 280
Qy 260 DHTGSMQHCQARQCAASH-----SPRMLPYENTNSENKPDPT----- 297
Db 281 RHRS-----SHRQDTQVSPVFPPTQKPVHQKPVLPFRPPEAGSPAADGPFH 327
Qy 298 -----KPDEKDVQWQNEWYIGEYSRQAVEDVLMKENKDGTFVLVDRDCSTKSKAEPVVLV 349
Db 328 SFPFNLTFADQEGELLGKPYAGACDKFAEALHRSNKDGSFLIRKSFHGDHSKQPYTLV 387
Qy 350 VFYGNKYVNVKIRFLESNQFALGTGLRGNEMFDSVEDIIETHTYTPILLIDGKD 404
Db 388 AFFNKRVINPVPIEATKQYALGKKNGEYFGSVVEIVNSHOHNPVLIDSQN 442
```

RESULT 11

```
US-09-513-999C-7233
; Sequence 7233, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7233
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 68
; OTHER INFORMATION: Xaa=Asp or Glu
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 69
; OTHER INFORMATION: Xaa=Ala or Asp or His or Ile or Asn or Pro or Ser or Thr or Tyr
```

```
US-09-513-999C-7233
Query Match 9.2%; Score 213; DB 2; Length 84;
Best Local Similarity 48.8%; Pred. No. 4.6e-12;
Matches 41; Conservative 17; Mismatches 26; Indels 0; Gaps 0;
Qy 1 MTSQGNKRTTKEGFDLRFQNVSLKRNRSWPSLSAKGRCAVLEPLPDHRRNLGAVPGG 60
Db 1 MNRQGNKRTTKESNDLKQNFSLPKNRSWPRINSATGQYQRMNKKLLDWFNFAAVLDG 60
Qy 61 EKCSNNNDYEDPEFQLLKAWPSMK 84
Db 61 AKGHSDDXDDPELRMEETWQSIK 84
```

RESULT 12

```
US-08-729-416C-1
; Sequence 1, Application US/08729416C
; Patent No. 6013767
; GENERAL INFORMATION:
```

```
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
; TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,416C
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, GLENN J.
; REGISTRATION NUMBER: 28458
; REFERENCE/DOCKET NUMBER: 7898/225948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-729-416C-1
```

```
Query Match 7.1%; Score 165.5; DB 2; Length 474;
Best Local Similarity 23.2%; Pred. No. 1.3e-06;
Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;
Qy 13 GFGLRFQNVSLKRNRSWPSLSAKGRCAVLEPLPDHRRNLGAVPGGKCSNNNDY--- 69
Db 94 GKSNIQFAGNISLTIISTASLNLRTDPSKQII--ANHMRISISFASGGDP--DTTDVAY 149
Qy 70 --EDP-----EFQLLKAWPSMKILPARIQESEVADTRYFQDMMEAPLLLPKASVSTERQ 123
Db 150 VAKDPVNRRAACHILECCDG---LAQDVIGSIGQAFELRFKQYLQCPKIP---ALHDMQ 203
Qy 124 TRDVRMTQLEEVDPKTFKDVRSQRFKGYTKINKTLPPLPPPAITLPPKKYQPLPPAPPE 183
Db 204 SLDEPWTE-EEGD-----GSDHPYNSIPSKMPPPGGFLDTLKLPRHAP-- 247
Qy 184 ESSAYFAPK-----PTFPE-VQRGPRQRSKDFSRVLGAEESHOTKPESSCP 231
Db 248 -DTAQFAGKEQTYVYQGRHLGDTFEGDWQQTFLRQSSD-----IVSTPEGLHVAPTGEAP 302
Qy 232 SSNQNTQKSPPAITASSYMPGKHSIQARDHTGSMQHCQAPQRCQAASHSPR-----MLPY 286
Db 303 -TYVNTQIQIPP-----QAWPA--AVSSAESSPKDLDFMKPF 336
Qy 287 ENTNSKPP-----DPTKPDEx---DVMQNEWYIGEYSRQAVEDVLMKE 326
Db 337 EDALKNQPLGVLVSKAASVECTSPVSPRAPDAKMLELQAEIYQEMSKAEAGLL--- 393
Qy 327 NKDGTFLVRDCSTKSKAEP--YVLVVFYGNKYVNVKIRFLESNQFALGTGLRGNEMFDS 384
Db 394 EKDGDFLVR---KSTTNPGSFVLTMGNHQAQKHLLLVDPE-----GTIRTKDRVDFS 442
Qy 385 VEDIIEHY--TYPEPILLIDGDKAARKKCYLTQPL 418
Db 443 ISHLINHLESLSPIV-----SAGSELC-LQOPV 470
```

```
RESULT 13
US-09-433-353-1
; Sequence 1, Application US/09433353
; Patent No. 6545141
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL, BRAIN -SPECIFIC ADAPTER MOLECULE, GENE THEREOF,
; AND ANTIBODY THERETO
; FILE REFERENCE: 7898/262241
; CURRENT APPLICATION NUMBER: US/09/433,353
; CURRENT FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 08/729,416
; PRIOR FILING DATE: 1996-10-11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-433-353-1

Query Match 7.1%; Score 165.5; DB 2; Length 474;
Best Local Similarity 23.2%; Pred. No. 1.3e-06;
Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;

QY 13 GFGDLRFQNVSLKRNWSPSLSSAKRCRAVLPLPDHRRNLGAVPGGKCNNDY--- 69
DB 94 GKSNIQFAGMSISLTITSTASLNLRTPDSKQII--ANHMRISIFASGSDP--DTTDYVAY 149
QY 70 --BDP-----EFQLLKAWPSMKILPARPIQSEVADTRYFODMMEAPLLLPKASVSTERQ 123
DB 150 VAKDPVNRACHILECCDG---LAQDVIGSIGQAFELRFKQYLCQCTKIP---ALHDMQ 203
QY 124 TRDVRMTQLEVDKPTFKDVRQRFGFKYTKINKTLPPLPPRAITLPPKQYQLPPAPPE 183
DB 204 SLDEPWE--EGD-----GSDHPYNSIPSKMPPPGGFLDTRLKPRHAP-- 247
QY 184 ESSAYFAPK-----PTPPE-VQRPRORSKDFSRVLGAEESHHQTPPESSCP 231
DB 248 -DTAQFAGKEQTYQGRHLGDTTFEGDWQQTPLRGSSD----IYSTPEGLKHVAPTGEAP 302
QY 232 SSNQNTQKSPATASSYMPGKHSIQARDHTGSMQHCPCQORCOAAASHSPR-----MLPY 286
DB 303 -TYVNTQOIIPP-----QAWPA--AVSSAESSPRKDLFDMKPPF 336
QY 287 ENTNSEKP-----DPTKPEK---DVMQNEWYIGYSRQAVEDVLMKE 326
DB 337 EDALKNQPLGPVLSKAASVEICISPVSPRAPDAKMLELQAEWTYQGEMSKAEGLL--- 393
QY 327 NKDGTFLVRCSTKSKAEP--YVLVVFYGNKVNKIRFLESNQOQFALGTGLRGNEFDS 384
DB 394 EKDGDFLVR---KSTTNPGSFVLTMHNGQAKHLLLVDP-----GTIRTKDRVDFS 442
QY 385 VEDIIEHY--TYPPILLIDGDKAARRKQCYLTQPL 418
DB 443 ISHLINHLESPLIV-----SAGSELCLQQPV 470

RESULT 14
US-08-729-416C-7
; Sequence 7, Application US/08729416C
; Patent No. 6013767
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
; AND ANTIBODY THEREOF, AND ANTIBODY THERETO
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA

; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,416C
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, GLENN J.
; REGISTRATION NUMBER: 28458
; REFERENCE/DOCKET NUMBER: 7898/225948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 594 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-729-416C-7

Query Match 7.1%; Score 165.5; DB 2; Length 594;
Best Local Similarity 23.2%; Pred. No. 1.1e-06;
Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;

QY 13 GFGDLRFQNVSLKRNWSPSLSSAKRCRAVLPLPDHRRNLGAVPGGKCNNDY--- 69
DB 214 GKSNIQFAGMSISLTITSTASLNLRTPDSKQII--ANHMRISIFASGSDP--DTTDYVAY 269
QY 70 --BDP-----EFQLLKAWPSMKILPARPIQSEVADTRYFODMMEAPLLLPKASVSTERQ 123
DB 270 VAKDPVNRACHILECCDG---LAQDVIGSIGQAFELRFKQYLCQCTKIP---ALHDMQ 323
QY 124 TRDVRMTQLEVDKPTFKDVRQRFGFKYTKINKTLPPLPPRAITLPPKQYQLPPAPPE 183
DB 324 SLDEPWE--EGD-----GSDHPYNSIPSKMPPPGGFLDTRLKPRHAP-- 367
QY 184 ESSAYFAPK-----PTPPE-VQRPRORSKDFSRVLGAEESHHQTPPESSCP 231
DB 368 -DTAQFAGKEQTYQGRHLGDTTFEGDWQQTPLRGSSD----IYSTPEGLKHVAPTGEAP 422
QY 232 SSNQNTQKSPATASSYMPGKHSIQARDHTGSMQHCPCQORCOAAASHSPR-----MLPY 286
DB 423 -TYVNTQOIIPP-----QAWPA--AVSSAESSPRKDLFDMKPPF 456
QY 287 ENTNSEKP-----DPTKPEK---DVMQNEWYIGYSRQAVEDVLMKE 326
DB 457 EDALKNQPLGPVLSKAASVEICISPVSPRAPDAKMLELQAEWTYQGEMSKAEGLL--- 513
QY 327 NKDGTFLVRCSTKSKAEP--YVLVVFYGNKVNKIRFLESNQOQFALGTGLRGNEFDS 384
DB 514 EKDGDFLVR---KSTTNPGSFVLTMHNGQAKHLLLVDP-----GTIRTKDRVDFS 562
QY 385 VEDIIEHY--TYPPILLIDGDKAARRKQCYLTQPL 418
DB 563 ISHLINHLESPLIV-----SAGSELCLQQPV 590

RESULT 15
US-09-433-353-7
; Sequence 7, Application US/09433353
; Patent No. 6545141
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL, BRAIN -SPECIFIC ADAPTER MOLECULE, GENE THEREOF,
; AND ANTIBODY THERETO
; FILE REFERENCE: 7898/262241
; CURRENT APPLICATION NUMBER: US/09/433,353
```

```
; CURRENT FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 08/729,416
; PRIOR FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-433-353-7

Query Match      7.4%  Score 165.5;  DB 2;  Length 594;
Best Local Similarity 23.2%;  Pred. No. 1.8e-06;
Matches 106;  Conservative 60;  Mismatches 161;  Indels 129;  Gaps 25;

Qy 13 GFGDLRFQNVSLKRNWPSLSAKGRCRAVLPLPDHRRNLGAVPGGKCNNDY--- 69
Db 214 GKSNLQFAGMSISLTISTASLNLRTPDSKQII--ANHHMSISFASGGDP--DTTDYVAY 269
Qy 70 --EDP-----FFQLLKAWPSMKILPARPIQSEYADTRYFQDMMEAPLLLPKASVSTERQ 123
Db 270 VAKDPVNRRAACHILECDG---LAQDVIGSIGQAFELRFKQYLQCP TKIP---ALHHRMQ 323
Qy 124 TRDVRMTQLEVDKPTFKVRSQRFKGVTKINKTLPPLPPPAITLPPKKYQPLPPAPPE 183
Db 324 SLDEPWTE-BEGD-----GSDHPYNSIPSKMPPPGGFLDTRLKPRPHAP-- 367
Qy 184 ESSAYFAPK-----PTPPE-VQGRQORSKDFSRVLGAEEESHHTKPESSCP 231
Db 368 -DTAQFAGKEQTYVQGRHLGDTTGEDWQQTFLRQSSD---IVSTPEGKLHVAPTGEAP 422
Qy 232 SSNQNTQKSPATASSYMGKHSIQARDHTGSMQHCPCAPORCQAAASHSPR-----MLPY 286
Db 423 -TYVNTQOIIP-----QAWPA--AVSSAESPRKDLFDMKPF 456
Qy 287 ENTNSEKP-----DPTKPEK---DVMQNEWYIGYSRQAVEDVLMKE 326
Db 457 EDALKNQPLGVLSKAASVFCISFVSPRAPDAKMLELQAEWTYQGEMSRKEAGLL--- 513
Qy 327 NKDGTFLVRDCSTKSKAEP--YVLVVFYGNKYNVNFLESNQQFALGTGLRGNEMFDS 384
Db 514 EKDGDFLVR---KSTTNPGSFVLTMHNGQAKHLLIVDPE-----GTIRTKDRVFD 562
Qy 385 VEDIIEHY--TYPEPILLIDGKDKAARKQCYLTPQL 418
Db 563 ISHLNHLESLEPIV-----SAGSELC-LOQPV 590
```

Search completed: March 22, 2006, 01:37:32
Job time : 47 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2006, 01:36:54 ; Search time 168 Seconds
(without alignments)
1081.880 Million cell updates/sec

Title: US-10-717-619-2
Perfect score: 2316
Sequence: 1 MTSQGNKRTTKEGFDLRFQ.....QPLPLARLLLTQYSSQALHHE 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	435	5	US-10-717-619-2
2	1316.5	56.8	428	3	US-09-966-955A-4
3	1316.5	56.8	428	4	US-10-104-047-3344
4	1316.5	56.8	443	3	US-09-966-955A-2
5	1306.5	56.4	428	5	US-10-717-619-4
6	1109.5	47.9	353	3	US-09-966-955A-6
7	398	17.2	533	5	US-10-723-860-2611
8	398	17.2	533	5	US-10-287-436A-431
9	398	17.2	533	5	US-10-287-436A-1132
10	319	13.8	456	5	US-10-719-993-671
11	309.5	13.4	433	5	US-10-719-993-672
12	258	11.1	128	5	US-10-617-316-224
13	212	9.2	46	3	US-09-864-761-38652
14	185.5	8.0	503	5	US-10-450-763-53213
15	177	7.6	67	4	US-10-639-067-114
16	161.5	7.0	344	5	US-10-483-506-21
17	159.5	6.9	812	5	US-10-732-923-13470
18	150.5	6.5	812	5	US-10-732-923-13476
19	143	6.2	848	4	US-10-322-281-371
20	143	6.2	881	4	US-10-322-281-374
21	140	6.0	540	5	US-10-723-860-4370
22	138.5	6.0	896	4	US-10-408-765A-1811
23	138.5	6.0	923	4	US-10-258-106-7
24	137	5.9	143	4	US-10-104-047-3542
25	134.5	5.8	715	4	US-10-435-696-98
26	134.5	5.8	3402	5	US-10-626-832-27
27	134	5.8	309	3	US-09-823-187-82

28	133	5.7	410	5	US-10-399-594-15	Sequence 15, Appl
29	133	5.7	675	3	US-09-977-260-4	Sequence 4, Appl
30	133	5.7	675	3	US-09-977-261-4	Sequence 4, Appl
31	133	5.7	675	4	US-10-186-399-3	Sequence 3, Appl
32	133	5.7	675	4	US-10-021-660-108	Sequence 108, App
33	133	5.7	675	4	US-10-211-462-129	Sequence 129, App
34	133	5.7	686	4	US-10-320-801-12	Sequence 12, Appl
35	133	5.7	697	4	US-10-353-690-34	Sequence 34, Appl
36	133	5.7	697	4	US-10-366-288-30	Sequence 30, Appl
37	133	5.7	697	5	US-10-509-599-22	Sequence 22, Appl
38	133	5.7	945	5	US-10-840-060-208	Sequence 208, App
39	133	5.7	945	6	US-11-097-143-25023	Sequence 25023, A
40	132.5	5.7	271	6	US-11-097-143-1605	Sequence 1605, Ap
41	132.5	5.7	592	4	US-10-369-493-1534	Sequence 1534, Ap
42	132.5	5.7	904	6	US-11-097-143-19503	Sequence 19503, A
43	130	5.6	551	4	US-10-369-493-13115	Sequence 13115, A
44	130	5.6	688	4	US-10-081-980B-1	Sequence 1, Appl
45	130	5.6	724	4	US-10-081-980B-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-10-717-619-2
; Sequence 2, Application US/10717619
; Publication No. US20040229237A1
; GENERAL INFORMATION:
; APPLICANT: GOITSUKA, RYO
; TITLE OF INVENTION: A mast cell-specific adapter molecules and cDNAs thereof
; FILE REFERENCE: 2003-1609/WMC/00653
; CURRENT APPLICATION NUMBER: US/10717, 619
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: PCT/JP00/06351
; PRIOR FILING DATE: 2000-9-17
; PRIOR APPLICATION NUMBER: JP11-263778
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-717-619-2

Query Match		100.0%	Score 2316;	DB 5;	Length 435;
Best Local Similarity		100.0%	Pred. No. 7.3e-169;		
Matches 435;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MTSQGNKRTTKEGFDLRFQNVSLLNKNSWPSSLSSAKGRCRAVLEPLPDHRRNLGAVPGG	60		
Db	1	MTSQGNKRTTKEGFDLRFQNVSLLNKNSWPSSLSSAKGRCRAVLEPLPDHRRNLGAVPGG	60		
Qy	61	EKCSNNVDYEDPEFQLLKAWPSNKIIPARPIQSEYADTRYFQDMMEAPLLLPKASVST	120		
Db	61	EKCSNNVDYEDPEFQLLKAWPSNKIIPARPIQSEYADTRYFQDMMEAPLLLPKASVST	120		
Qy	121	EROTRDRVMTQLEVDKPTFKDVRQRFKGYKTKTKNTPLPPRPAITLPLKKYQPLPPA	180		
Db	121	EROTRDRVMTQLEVDKPTFKDVRQRFKGYKTKTKNTPLPPRPAITLPLKKYQPLPPA	180		
Qy	181	PPESAYFAPKPTTFPEVQGRQPSAKDFSRVLGAEEESHHTQKPESSCPSSNQNTQKS	240		
Db	181	PPESAYFAPKPTTFPEVQGRQPSAKDFSRVLGAEEESHHTQKPESSCPSSNQNTQKS	240		
Qy	241	PPAIASSYMPGKHSIQARDHTGSMQHCPCQAORCAAAASHSPRLPYENTNSEKPDTPKPD	300		
Db	241	PPAIASSYMPGKHSIQARDHTGSMQHCPCQAORCAAAASHSPRLPYENTNSEKPDTPKPD	300		
Qy	301	EKDVQNEWYIGYSYRQAVEDVLMEKNGKDTFLVRDCSTKSKABPYVLVVFYGNKYNVVK	360		
Db	301	EKDVQNEWYIGYSYRQAVEDVLMEKNGKDTFLVRDCSTKSKABPYVLVVFYGNKYNVVK	360		

```
Qy 361 IRPLESNOQFALGTGLRGNEFDSVEDIIHYTYFPILLIDGDKDAARRKOCYLTQPL 420
Db 361 IRPLESNOQFALGTGLRGNEFDSVEDIIHYTYFPILLIDGDKDAARRKOCYLTQPL 420
Qy 421 ARLLLTQYSSQALHE 435
Db 421 ARLLLTQYSSQALHE 435

RESULT 2
US-09-966-955A-4
; Sequence 4, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gena S.
; APPLICANT: Kanner, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-length Human
; TITLE OF INVENTION: Clink-related Gene, MIST (Mast Cell Immunoreceptor)
; TITLE OF INVENTION: Signal Transducer)
; FILE REFERENCE: 3053-4113US1
; CURRENT APPLICATION NUMBER: US/09/966.955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 428
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN MIST SPLICE VARIANT CLONE #7, AMINO ACID
; OTHER INFORMATION: SEQUENCE
US-09-966-955A-4

Query Match 56.8%; Score 1316.5; DB 3; Length 428;
Best Local Similarity 62.1%; Pred. No. 2.6e-92;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;

Qy 1 MTSQGNKRTTKEGFDLRFQNVSLLKXRSWPSLSAKGRCAVLEPLPDHRRNLAVPGG 60
Db 1 MNRQGNKRTTKEGSNDLKQNFSLPKNRSWPRINSATGQYQRMNKPLLDWERNFAAVLDG 60
Qy 61 EKCSNNDYEDPFPQLLKAWPSMKILPARPIQSEYADTRYFODMMEAPLLLPKASVST 120
Db 61 AKGHSDDDDYDDELRLMEETWQSILKILPARPIKESEYADTHYFKVAMDTPLDLTRTSISI 120
Qy 121 ERQTRDVRMTQLEEVDPKTFKDVRSQRFKGYTKINKTLPPLPPPAITLPPKYQPLPPA 180
Db 121 GQPTWNTQ-TRLERVDKPIKDVRSQNIKGDASVRKNKIPLPPRPLITLPPKYQPLPP- 178
Qy 181 PPRESSAYFAPKPTFPFVQGRPRQSAKDFSRVLGAEESHQTKPESSCPSSNQNTQKS 240
Db 179 EPSSRPPLSQRTTFPEVQRMPSQISLRDLSEVLEAEKVPHNQRPKPESTHLENQNTQEI 238
Qy 241 PDEKDVQNEWYIGEYSRQAVEAFMKNKDGSLVRDCSTKSKAEPYVLVVFYGNKVYN 358
Db 239 PLAISSSFTTSHSVQNRDHRGMQPCSPQRCQPASPSCSPHENILPKYKTSWRPFPK 298
Qy 359 VKIRFLESNOQFALGTGLRGNEFDSVEDIIHYTYFPILLIDGDK-AARRKOCYLTQ 417
Db 359 VKIRFLERNQOQFALGTGLRGDEKFDSDVEDIIHYKNFPILLIDGKDKTGVRHKOCHLTQ 418
Qy 418 LPLRLALL 425
Db 419 LPLTRHLL 426

RESULT 4
US-09-966-955A-2
; Sequence 2, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gena S.
; APPLICANT: Kanner, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-length Human
; TITLE OF INVENTION: Clink-related Gene, MIST (Mast Cell Immunoreceptor)
```

```
Db 419 LPLTRHLL 426

RESULT 3
US-10-104-047-3344
; Sequence 3344, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3344
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-104-047-3344

Query Match 56.8%; Score 1316.5; DB 4; Length 428;
Best Local Similarity 62.1%; Pred. No. 2.6e-92;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;

Qy 1 MTSQGNKRTTKEGFDLRFQNVSLLKXRSWPSLSAKGRCAVLEPLPDHRRNLAVPGG 60
Db 1 MNRQGNKRTTKEGSNDLKQNFSLPKNRSWPRINSATGQYQRMNKPLLDWERNFAAVLDG 60
Qy 61 EKCSNNDYEDPFPQLLKAWPSMKILPARPIQSEYADTRYFODMMEAPLLLPKASVST 120
Db 61 AKGHSDDDDYDDELRLMEETWQSILKILPARPIKESEYADTHYFKVAMDTPLDLTRTSISI 120
Qy 121 ERQTRDVRMTQLEEVDPKTFKDVRSQRFKGYTKINKTLPPLPPPAITLPPKYQPLPPA 180
Db 121 GQPTWNTQ-TRLERVDKPIKDVRSQNIKGDASVRKNKIPLPPRPLITLPPKYQPLPP- 178
Qy 181 PPRESSAYFAPKPTFPFVQGRPRQSAKDFSRVLGAEESHQTKPESSCPSSNQNTQKS 240
Db 179 EPSSRPPLSQRTTFPEVQRMPSQISLRDLSEVLEAEKVPHNQRPKPESTHLENQNTQEI 238
Qy 241 PPAIASSSYMPGKHISIQARDHTGSMQHCPAQRQAAASHSP--RMLPYENTNSEKDPDK 298
Db 239 PLAISSSFTTSHSVQNRDHRGMQPCSPQRCQPASPSCSPHENILPKYKTSWRPFPK 298
Qy 299 PDEKDVQNEWYIGEYSRQAVEAFMKNKDGSLVRDCSTKSKAEPYVLVVFYGNKVYN 358
Db 299 SDRKDQVHNEWYIGEYSRQAVEAFMKNKDGSLVRDCSTKSKAEPYVLVAVYENKVYN 358
Qy 359 VKIRFLESNOQFALGTGLRGNEFDSVEDIIHYTYFPILLIDGDK-AARRKOCYLTQ 417
Db 359 VKIRFLERNQOQFALGTGLRGDEKFDSDVEDIIHYKNFPILLIDGKDKTGVRHKOCHLTQ 418
Qy 418 LPLRLALL 425
Db 419 LPLTRHLL 426

RESULT 4
US-09-966-955A-2
; Sequence 2, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gena S.
; APPLICANT: Kanner, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-length Human
; TITLE OF INVENTION: Clink-related Gene, MIST (Mast Cell Immunoreceptor)
```

```

; TITLE OF INVENTION: Signal Transducer)
; FILE REFERENCE: 3053-4113US1
; CURRENT APPLICATION NUMBER: US/09/966,955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN FULL-LENGTH MIST cDNA CLONE #8, TRANSLATED
; OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-2

Query Match          56.8%; Score 1316.5; DB 3; Length 443;
Best Local Similarity 62.1%; Pred. No. 2.8e-92;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;

Qy 1 MTSQGNKRTTKEGFGDLRFQNVSLLNKRSWPSLSSAKGRCAVLEPLPDHRRNLGVPGG 60
Db 16 MNRQGNKRTTKEGSDNLKFNQFSLPKNRSWPRINSATGQYQRMNKLPLDWRNFAA 75
Qy 61 EKNSNNDDYEDPEFOLLKAWPSMKILPARPIQSEYADTRYFQDMMEAPLLLPKASVST 120
Db 76 AKGHSDDDDYDDPELRMEETWQSIKILPARPIKESEYADTRYFKVAMDTPLDTRTSIS 135
Qy 121 ERQTRDVRMTQLEEVDPKTFKQVRSQRFKGYTKINKTLPPLPPRAITLIPKYYQPLPP 180
Db 136 GQPTWNTQ-TRLERVDKPISRDLSVLEAEKVPHNQKPESTHLENQNTQEI 193
Qy 181 PPESSAYFAPKPTFFEVQGRPRORSKADFSRVLGAEEESHQTKPESSCPSSNQNTQKS 240
Db 194 EPSSRRPPLSQRHTFFEVQMPQSIQLRDLSEVLEAEKVPHNQKPESTHLENQNTQEI 253
Qy 241 PPAIASSSYMPGKHISIQARDHTGSMOHCPCPAORCOAAASHSP--RMLPYENTNSEKPDPTK 298
Db 254 PLAISSSFTTSNHSVQNRDHRGQMPCSPQRCQPASPSCSPHENILPYKTSWRPFPKR 313
Qy 299 PDEKDVQWNEWYIGEYSRQAVEDVLMEKNKGDTFLVRDCSTKSKAEYPVLVVFYGNKVYN 358
Db 314 SDRKDVQWNEWYIGEYSRQAVEAEAFMEKNKGDTFLVRDCSTKSKAEYPVLVVFYGNKVYN 373
Qy 359 VKIRFLRNOQFALGTGLRGNEFDSVEDIIHYHTYTPILLIDGKDK-AARRKQCYLTQP 417
Db 374 VKIRFLRNOQFALGTGLRGDEKFDSDVEDIIHYKNFPILLIDGKDKTGVHRKQCHLTQP 433
Qy 418 LPLARLLL 425
Db 434 LPLTRHLL 441

RESULT 5
US-10-717-619-4
; Sequence 4, Application US/10717619
; Publication No. US20040229237A1
; GENERAL INFORMATION:
; APPLICANT: GOITSUKA, Ryo
; TITLE OF INVENTION: A mast cell-specific adapter molecules and cDNAs thereof
; FILE REFERENCE: 2003-1609/WMC/00653
; CURRENT APPLICATION NUMBER: US/10/717,619
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: PCT/JP00/06351
; PRIOR FILING DATE: 2000-9-17
; PRIOR APPLICATION NUMBER: JP11-263778
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 428
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
US-10-717-619-4

Query Match          56.4%; Score 1306.5; DB 5; Length 428;
Best Local Similarity 61.7%; Pred. No. 1.5e-91;
Matches 264; Conservative 47; Mismatches 112; Indels 5; Gaps 4;

Qy 1 MTSQGNKRTTKEGFGDLRFQNVSLLNKRSWPSLSSAKGRCAVLEPLPDHRRNLGVPGG 60
Db 1 MNRQGNKRTTKEGSDNLKFNQFSLPKNRSWPRINSATGQYQRMNKLPLDWRNFAA 60
Qy 61 EKNSNNDDYEDPEFOLLKAWPSMKILPARPIQSEYADTRYFQDMMEAPLLLPKASVST 120
Db 61 AKGHSDDDDYDDPELRMEETWQSIKILPARPIKESEYADTRYFKVAMDTPLDTRTSIS 120
Qy 121 ERQTRDVRMTQLEEVDPKTFKQVRSQRFKGYTKINKTLPPLPPRAITLIPKYYQPLPP 180
Db 121 GQPTWNTQ-TRLERVDKPISRDLSVLEAEKVPHNQKPESTHLENQNTQEI 178
Qy 181 PPESSAYFAPKPTFFEVQGRPRORSKADFSRVLGAEEESHQTKPESSCPSSNQNTQKS 240
Db 179 EPSSRRPPLSQRHTFFEVQMPQSIQLRDLSEVLEAEKVPHNQKPESTHLENQNTQEI 238
Qy 241 PPAIASSSYMPGKHISIQARDHTGSMOHCPCPAORCOAAASHSP--RMLPYENTNSEKPDPTK 298
Db 239 PLAISSSFTTSNHSVQNRDHRGQMPCSPQRCQPASPSCSPHENILPYKTSWRPFPKR 298
Qy 299 PDEKDVQWNEWYIGEYSRQAVEDVLMEKNKGDTFLVRDCSTKSKAEYPVLVVFYGNKVYN 358
Db 299 SDRKDVQWNEWYIGEYSRQAVEAEAFMEKNKGDTFLVRDCSTKSKAEYPVLVVFYGNKVYN 358
Qy 359 VKIRFLRNOQFALGTGLRGNEFDSVEDIIHYHTYTPILLIDGKDK-AARRKQCYLTQP 417
Db 359 VKIRFLRNOQFALGTGLRGDEKFDSDVEDIIHYKNFPILLIDGKDKTGVHRKQCHLTQP 418
Qy 418 LPLARLLL 425
Db 419 LPLTRHLL 426

RESULT 6
US-09-966-955A-6
; Sequence 6, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gena S.
; APPLICANT: Kanner, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-length Human
; TITLE OF INVENTION: Clnk-related Gene, MIST (Mast Cell Immunoreceptor
; TITLE OF INVENTION: Signal Transducer)
; FILE REFERENCE: 3053-4113US1
; CURRENT APPLICATION NUMBER: US/09/966,955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 353
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN MIST SPICE VARIANT CLONE #12, TRANSLATED
; OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-6

Query Match          47.9%; Score 1109.5; DB 3; Length 353;
Best Local Similarity 64.4%; Pred. No. 1.5e-76;
Matches 226; Conservative 31; Mismatches 89; Indels 5; Gaps 4;
```


This Page Blank (Uplo)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 22, 2006, 01:37:44 ; Search time 23 Seconds
(without alignments)
541.346 Million cell updates/sec

Title: US-10-717-619-2

Perfect score: 2316

Sequence: 1 MTSQGNKRTTKEGFDLRQ.....QPLPLARLLLTQYSSQALHE 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 169630 seqs, 28622889 residues

Total number of hits satisfying chosen parameters: 169630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1316.5	56.8	428	7	US-11-072-512-3344
2	138.5	6.0	950	6	US-10-501-035-357
3	137	5.9	143	7	US-11-072-512-3542
4	136.5	5.9	561	6	US-10-467-033-2
5	136.5	5.9	561	6	US-10-467-033-6
6	128.5	5.5	393	7	US-11-221-263-26
7	128	5.5	724	7	US-11-109-156-19
8	127.5	5.5	863	7	US-11-169-041-167
9	121.5	5.2	2344	6	US-10-330-773-627
10	117	5.1	428	7	US-11-221-263-24
11	116.5	5.0	536	6	US-10-821-231C-1
12	116	5.0	667	6	US-10-821-234-1477
13	116	5.0	667	7	US-11-124-368A-250
14	116	5.0	686	7	US-11-124-368A-239
15	116	5.0	708	7	US-11-072-175-158
16	116	5.0	726	7	US-11-124-368A-247
17	116	5.0	722	7	US-11-124-368A-248
18	116	5.0	750	7	US-11-124-368A-244
19	116	5.0	750	7	US-11-124-368A-249
20	116	5.0	777	7	US-11-124-368A-238
21	116	5.0	778	7	US-11-124-368A-240
22	116	5.0	791	7	US-11-124-368A-245
23	115.5	5.0	565	6	US-10-915-161-2
24	115	5.0	2801	6	US-10-330-773-630
25	114.5	4.9	533	7	US-11-230-995-3

26	114	4.9	460	7	US-11-087-099-9781	Sequence 9781, Ap
27	112.5	4.9	798	6	US-10-514-531-2	Sequence 2, Appli
28	112.5	4.9	941	7	US-11-169-232-14	Sequence 14, Appl
29	112.5	4.9	1022	7	US-11-169-232-84	Sequence 84, Appl
30	112.5	4.9	1038	7	US-11-169-232-74	Sequence 74, Appl
31	112.5	4.9	1049	7	US-11-169-232-58	Sequence 58, Appl
32	112.5	4.9	1140	7	US-11-169-232-104	Sequence 104, Appl
33	112.5	4.9	1270	7	US-11-169-232-44	Sequence 44, Appl
34	112.5	4.9	1311	7	US-11-169-232-42	Sequence 42, Appl
35	112.5	4.9	1313	7	US-11-169-232-142	Sequence 142, App
36	112.5	4.9	1314	7	US-11-169-232-50	Sequence 50, Appl
37	112.5	4.9	1320	7	US-11-169-232-46	Sequence 46, Appl
38	112.5	4.9	1320	7	US-11-169-232-60	Sequence 60, Appl
39	112.5	4.9	1354	7	US-11-169-232-48	Sequence 48, Appl
40	112.5	4.9	1361	7	US-11-169-232-40	Sequence 40, Appl
41	112.5	4.9	1363	7	US-11-169-232-52	Sequence 52, Appl
42	112.5	4.9	1404	7	US-11-169-232-2	Sequence 2, Appli
43	112.5	4.9	1404	7	US-11-169-232-62	Sequence 62, Appl
44	111.5	4.8	410	7	US-11-096-568A-22820	Sequence 22820, A
45	111.5	4.8	428	7	US-11-096-568A-22819	Sequence 22819, A

ALIGNMENTS

RESULT 1

US-11-072-512-3344
; Sequence 3344, Application US/11072512
; Publication No. US20060029945A1

GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3344
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3344

Query Match 56.8%; Score 1316.5; DB 7; Length 428;

Best Local Similarity 62.1%; Pred. No. 2.8e-100; Indels 5; Gaps 4;
Matches 266; Conservative 46; Mismatches 111;

Qy 1 MTSQGNKRTTKEGFDLRQFNVSLLKNRSWPSLSAKGRCAVLPLPDHRRNLGYPGG 60

Db 1 MNEQGNKRTTKEGNDLKFQNFSLPKNRSWPRINSATQYQRMNKPDLDERNFAAVLDG 60

Qy 61 EKCNSNDYEDPDEPQLLKAWPSMKILLPARIQSEYADTRYFQDMMEAPLLLPKASVST 120

```
Db 61 AKGSDDDYDDPELRMEETWQSIKILPARIKESYADTHYFKVAMDTPLPLTRTSISI 120
QY 121 ERQTRDVRMTQLEEVDRKPTFKDVRQRKFKYTKINKTLPPLPPPAITLPKKYQPLPPA 180
Db 121 GQPTWNTQ-TRLERVDKPIKSDVRQRKFKYTKINKTLPPLPPPAITLPKKYQPLPP- 178
QY 181 PPESSAYFAPKPTFFVQVRGPRQRSKDFSRVLGAEEESHQTKPESSCPSSNONTOKS 240
Db 179 EPSSRPPLSRHTFFVQVRMPQSIQSLRDLSEVLEAEKVPKPNQRKPESTHLLNQNQTQEI 238
QY 241 PPAIASSYMPGKHISQIARDHTSMQHCPCAPQRCQAASHSP--RMLPYENTNSEKDPPTK 298
Db 239 PLAISSTFTSNHSVQNRDHRGMQPCSPQRCPPASCPHENILPKYKTSWRPFPKR 298
QY 299 PDEKVMQNEWYIGEYSRQAVEDVLMKENKDGTFVLVRDCSTKSKABPYVLVVFYGNKVYN 358
Db 299 SDRKDVQHNNEWYIGEYSRQAVEAFMKNKDGSLVRDCSTKSKABPYVLVVFYGNKVYN 358
QY 359 VKIRFLESNOQFALGTGLRGNEPDSVEDIIEHYTFPILLIDGKDK-AARRKQCYLTQP 417
Db 359 VKIRFLENOQFALGTGLRGDEKFDVEDIIEHYKFNFPILLIDGKDKGVHRRKQCHLTQP 418
QY 418 LPLARLLL 425
Db 419 LPLTRHLL 426

RESULT 2
US-10-501-035-357
; Sequence 357, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 357
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-501-035-357

Query Match 6.0%; Score 138.5; DB 6; Length 950;
Best Local Similarity 22.8%; Pred. No. 0.0023;
Matches 80; Conservative 41; Mismatches 153; Indels 77; Gaps 15;

QY 50 HRENLAGVPG-----GEKCNNDYDEPEFO-LLKAMP-----SMKI 85
Db 625 HLQNLARTKGLILENKLVLLEYIGPEADPNLAPAPQLQDLASSCPQEVSVQOQESVST 694
QY 86 LPAR---PIQESYADTRYFQDMWEAPLLL-----PPKASVSTERQTRDVRM 129
Db 685 LPASVHPQLSPROSLETQYLQHRLOKPSLLSKAQTLCQLYCKEPPR---SLEQQLQEHRL 741
QY 130 TQLEEVDRKPTFKDVRQRKFKYTKINKTLPPLPPPAITLPKKYQPLPPAPESSAYF 189
Db 742 QQ-----KRLFLQKSOLOQAYFNQMQIAESSYPQSQQLPLPR--QETPP-PSQQAAPPFS 793
QY 190 AKPTPEVQGRQRSAKDFSRVLGAEEESHQTKPESS-----CPSSNONTOKSP 242
Db 794 LTQPLSLVLEPSSEQ---MQYSFPLSQYQEMQIQLPSTSGPRAAPPLPQLQOQOQPPPP 850
QY 243 AIASSSYMPGKHISQIARDHTSMQHCPCAPQRCQAASHSPRMLPYENTNSEKPDPTKPEK 302
Db 851 PPPPPPPQCGAAPLQ---FSYQTCLPASAAPADPYPTPCQYPVDGAQQSLLTGPD-- 905
```

```
QY 303 DVNQNEWYIGEYSRQAVEDVLMKENKDGTFVLVRDCSTKSKAEP-----YVLV 349
Db 906 ----CPRSPGLQBPSSYDFLALSELPLGLF---DCEMLDAVDPOHNGYVLV 949

RESULT 3
US-11-072-512-3542
; Sequence 3542, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAWATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3542
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3542

Query Match 5.9%; Score 137; DB 7; Length 143;
Best Local Similarity 25.5%; Pred. No. 0.00027;
Matches 39; Conservative 23; Mismatches 57; Indels 34; Gaps 5;

QY 285 PYENTNSEKPDPTKPEKDVQWQNEWYIGEYSRQAVEDVLMKENK----- 328
Db 5 PLPRTSVVPRPTTAQES-----TRNGTADAASKGKSGSLPSVAPTGSASAA 51
QY 329 -DQTFVLVRDCSTKSKABPYVLVVFYGNKVNVKIRFLESNQOQFALG-TGLRGNEPDSVE 386
Db 52 EDGAYTVRSGSGPHGQPTFLAVLLRGRVNTFIRLDGGRHYALGREGNRBELFSSVA 111
QY 387 DIIHYTYFPILLIDGDKAARRKQCYL--TQP 417
Db 112 AMQVHFMMHPLPLVD-RHSGSRELTCLLPFTKP 143

RESULT 4
US-10-467-033-2
; Sequence 2, Application US/10467033
; Publication No. US20060019248A1
; GENERAL INFORMATION:
; APPLICANT: Tiziani, Valdenize
; APPLICANT: Reichenberger, Ernest
; APPLICANT: Ueki, Yasuyoshi
; APPLICANT: Olsen, Bjorn R.
; TITLE OF INVENTION: Mutant SH3-Binding Protein Compositions and Methods
; FILE REFERENCE: H0498.70204US00
```

```
; CURRENT APPLICATION NUMBER: US/10/467,033
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,129
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-467-033-2

Query Match          5.9%; Score 136.5; DB 6; Length 561;
Best Local Similarity 22.4%; Pred. No. 0.0017;
Matches 94; Conservative 44; Mismatches 138; Indels 143; Gaps 23;

QY 66 NNDYEDPEFOLLKAWPSMKILPARPIQSEYADT-----RYFQDMMEAPLLLPKASVS 119
DB 171 NEDYEHDD-----EDDSYLEPDSPEGRLEDAIMHPPAYPPP--PVP 210
QY 120 TERQTRDVRMTQLEVDKPTFKDVRQRFKGYTKINKTP-LPPPPPAITL----- 170
DB 211 TPR-----KPAFSDMP-----RAHSFTSGPGPLPPPPPKHGLPDVGLAAE 252
QY 171 -----PKYQPLP--PAPPEE-SSAYFAPKPTTFEVORGRQR----- 205
DB 253 DSKRDPLCPRAEPCRPVATPRMSDPLSTMTPTAGLRKPCFCFRESASPSPEPWTGPH 312
QY 206 -SAKDFSRVLGABEESHQTK-----PSSSCPSSNQ-----TQKSPPAIAS 246
DB 313 GACSTSSAATMATATSRNCDKLSFHLSPRGPTTSEPPVPANKPKFLKIAEDPPREAA 372
QY 247 SSTMPGKHSTQARDHTGSMQHCPCAQRCQAAASHSPR-----MLP-----YENTNS 291
DB 373 ---MPGLFVPPVAPRPPALK-LPVPPEAMARPAVLPRPEKQPLHLQRSPPDGQSFSP 428
QY 292 EKP-DPTKPD-----EKDVQWNEWYIGYSRQAVEDVLMK-----ENKDGTFLVR 335
DB 429 EKPRQPSQADTGDDSDDEYKVP LPSNVFVNTTESCEVERLFKATSPRGEQDGLYCIR 488
QY 336 DCSTKSKAEPYVLVWF--YGNKVNVKIRFLESNQOQFALGTGLRGNEMFDSVEDIIEHY 392
DB 489 NSSTKSGK---VLVWDETSNKNRYRI--FEKDSKFY-----LEGEVLFVSGSMVEHY 538

RESULT 5
US-10-467-033-6
; Sequence 6, Application US/10467033
; Publication No. US20060019248A1
; GENERAL INFORMATION:
; APPLICANT: Tiziani, Valdenize
; APPLICANT: Reichenberger, Ernest
; APPLICANT: Ueki, Yasuyoshi
; APPLICANT: Olsen, Bjorn R.
; TITLE OF INVENTION: Mutant SH3-Binding Protein Compositions and Methods
; FILE REFERENCE: H0498.70204U00
; CURRENT APPLICATION NUMBER: US/10/467,033
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,129
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 561
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (415)..(415)
; OTHER INFORMATION: Xaa = any amino acid
; OTHER INFORMATION: provided that the mutant domain is not wild type
; FEATURE:
; NAME/KEY: MISC_FEATURE

; CURRENT APPLICATION NUMBER: US/10/467,033
; CURRENT FILING DATE: 2003-08-01
; PRIOR APPLICATION NUMBER: US 60/266,129
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-467-033-6

; LOCATION: (418)..(418)
; OTHER INFORMATION: Xaa = any amino acid
; OTHER INFORMATION: provided that the mutant domain is not wild type
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (420)..(420)
; OTHER INFORMATION: Xaa = any amino acid
; OTHER INFORMATION: provided that the mutant domain is not wild type
US-10-467-033-6

Query Match          5.9%; Score 136.5; DB 6; Length 561;
Best Local Similarity 22.4%; Pred. No. 0.0017;
Matches 94; Conservative 44; Mismatches 138; Indels 143; Gaps 23;

QY 66 NNDYEDPEFOLLKAWPSMKILPARPIQSEYADT-----RYFQDMMEAPLLLPKASVS 119
DB 171 NEDYEHDD-----EDDSYLEPDSPEGRLEDAIMHPPAYPPP--PVP 210
QY 120 TERQTRDVRMTQLEVDKPTFKDVRQRFKGYTKINKTP-LPPPPPAITL----- 170
DB 211 TPR-----KPAFSDMP-----RAHSFTSGPGPLPPPPPKHGLPDVGLAAE 252
QY 171 -----PKYQPLP--PAPPEE-SSAYFAPKPTTFEVORGRQR----- 205
DB 253 DSKRDPLCPRAEPCRPVATPRMSDPLSTMTPTAGLRKPCFCFRESASPSPEPWTGPH 312
QY 206 -SAKDFSRVLGABEESHQTK-----PSSSCPSSNQ-----TQKSPPAIAS 246
DB 313 GACSTSSAATMATATSRNCDKLSFHLSPRGPTTSEPPVPANKPKFLKIAEDPPREAA 372
QY 247 SSTMPGKHSTQARDHTGSMQHCPCAQRCQAAASHSPR-----MLP-----YENTNS 291
DB 373 ---MPGLFVPPVAPRPPALK-LPVPPEAMARPAVLPRPEKQPLHLQRSPPDGQSFSP 428
QY 292 EKP-DPTKPD-----EKDVQWNEWYIGYSRQAVEDVLMK-----ENKDGTFLVR 335
DB 429 EKPRQPSQADTGDDSDDEYKVP LPSNVFVNTTESCEVERLFKATSPRGEQDGLYCIR 488
QY 336 DCSTKSKAEPYVLVWF--YGNKVNVKIRFLESNQOQFALGTGLRGNEMFDSVEDIIEHY 392
DB 489 NSSTKSGK---VLVWDETSNKNRYRI--FEKDSKFY-----LEGEVLFVSGSMVEHY 538

RESULT 6
US-11-221-263-26
; Sequence 26, Application US/11221263
; Publication No. US20060051823A1
; GENERAL INFORMATION:
; APPLICANT: The United States of America as Represented by the Secretary of the
; APPLICANT: Department of Health and Human Services, Centers for Disease Control and
; APPLICANT: Prevention
; APPLICANT: Liu, Hsi
; APPLICANT: Steiner, Bret
; APPLICANT: Berta, Rodes
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING TREPONEMA PALLIDUM
; FILE REFERENCE: 6395-61666
; CURRENT APPLICATION NUMBER: US/11/221,263
; CURRENT FILING DATE: 2005-09-06
; PRIOR APPLICATION NUMBER: US/10/017,168
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: PCT/US00/16425
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/138,981
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Treponema pallidum
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Subspecies: pertenu (CDC-1 strain)
```

US-11-221-263-26

Query Match 5.5%; Score 128.5; DB 7; Length 393;
Best Local Similarity 21.3%; Pred. No. 0.0048;
Matches 78; Conservative 41; Mismatches 154; Indels 93; Gaps 14;
QY 52 RNLGAVGGGKSNNDYDEPEFOLLKAWPSMKILPARPIQES-----EYADT 99
DB 63 RTLGTVRGGSTQDQ-----LSLASLPSR--VPAAPQARDPLSSPPAGHTVPEYRDT 113
QY 100 RYFQDMMEAPLLPPKASVSTERQTRDVRMTQLEEVNDKPTFKDVRQSRQKGFYKTKINT 159
DB 114 VIFDD---PRLVSPLSRGEGEREVDV-----PKVVPASEREGC----- 150
QY 160 PLPPPRPAIPLPKYQPLPAPPESSAYFAPKPTFPFVQVGRQPRQSAKDFSRVL----- 214
DB 151 ----EREVDVPKVVPASEREGGEREVEDVPKVVPASEREGGEREVEDVPKVVPASE 206
QY 215 --GAEEESHQTK--PESSCPSSNQNTQKSPPAIASSSYMPGKHSIQARDHTGSMQHCP 269
DB 207 REGGEREVEDVPKVVPASEREGGEREVEDVPKVVPASEREGGEREVASQHTKQPSH-- 264
QY 270 AQCQAAASHSPRLPYENTNSKPDPTKPD--EKDVWQNEWYIGYSRQAVEDVLMKEN 327
DB 265 -----SVSNAPNQ--FRNPEGELPF-TLPDLSESEIIVVEEQKGRAHPQVIEGAPRGL 316
QY 328 KQGTFLVDCSTKSKAEPPYVLVVFYG-----NKVYNVKIRP-----L 364
DB 317 QPGEYYVQIAVFHDAIQVQSIHVHYGYEYPIAVEQDIHEGKVRFTVCVGPVQKDERGAVL 376
QY 365 ESNQOF 370
DB 377 ENQRF 382

RESULT 7

US-11-109-156-19
; Sequence 19, Application US/11109156
; Publication No. US20050250144A1
; GENERAL INFORMATION:
; APPLICANT: Toshio Ota
; APPLICANT: Takao Isogai
; APPLICANT: Tetsuo Nishikawa
; APPLICANT: Koji Hayashi
; APPLICANT: Kaoru Otsuka
; APPLICANT: Jun-ichi Yamamoto
; APPLICANT: Shizuko Ishii
; APPLICANT: Tomoyasu Sugiyama
; APPLICANT: Ai Makamatsu
; APPLICANT: Keiichi Nagai
; APPLICANT: Tetsuji Otsuki
; APPLICANT: Shin-ichi Funahashi
; APPLICANT: Chiaki Senoo
; APPLICANT: Jun-ichi Nezu
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN KINASE/PROTEIN
; FILE REFERENCE: PHOSPHATASE
; CURRENT APPLICATION NUMBER: US/11/109,156
; CURRENT FILING DATE: 2005-04-19
; PRIOR APPLICATION NUMBER: US/10/060,065
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05061
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02

; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-109-156-19

Query Match 5.5%; Score 128; DB 7; Length 724;
Best Local Similarity 20.6%; Pred. No. 0.012;
Matches 78; Conservative 45; Mismatches 135; Indels 120; Gaps 14;
QY 81 PSMKILPARPIQSEVADTRYFQDMMEAPLLLPKAS----- 117
DB 85 PTPKPRPRPLVPAP--GSSKTEADVEQALTLPLDLAEQFAPPDIAPPLLIKLVAEIEKKG 143
QY 118 -----VSTERQTRDVRMTQLEEVNDKPTFK-----DVRQSRQKGFYKTKINTPLPPR 165
DB 144 LECSTLYRTQSSNLNLAELRQLDCLDTPFSDVLEMDVHVLADAFKRYL-----LDLPN 195
QY 166 PAITLPKKYQPLPAPPESSAYFAP-----KPTFPEVQVGRQPRQSAKDFSRVLGAEE 218
DB 196 PVTAPAAVYSEMISLAPEVQSSEYIQLLKXLIIRSPISPHQYWLTLQVLLKHFFKL----- 250
QY 219 ESHHQTKESSCPSSNQNTQKSPPAIASSSYMPGKHSIQARDHTGSM----- 265
DB 251 -----SQTSSKNLLNARVLSEIFSPMLFRFSAASSDNTENLIKVIEILISTEMN 299
QY 266 --QHCFAQRCQAAASHSPRLPYENTNSKPDPTK-----DEKDVWQNEWYIGY 314
DB 300 EQPAPA-----LP-----PKPPKPTTVANNQNNMNSLQNAEWYWGDI 338
QY 315 SROAVEDVLMKENKQGTFLVDCSTKSKAEPPYVLVVFYGNKYNVKIRPLESNOQFALGT 374
DB 339 SREEVNEKL-RDTADGTFLVRDASTKMHGD-YTLIRKGN--NKLKIFHRDQKYGFS 394
QY 375 GLRGNEFMFDSVEDIIEHY 392
DB 395 PL---TFSSVVELINHY 408

RESULT 8

US-11-169-041-167
; Sequence 167, Application US/11169041
; Publication No. US20060019284A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF
; TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE
; TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 10001 NP
; CURRENT APPLICATION NUMBER: US/11/169,041
; CURRENT FILING DATE: 2005-06-28
; PRIOR APPLICATION NUMBER: 60/584,405
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 527
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 167
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-169-041-167

Query Match 5.5%; Score 127.5; DB 7; Length 863;
Best Local Similarity 22.7%; Pred. No. 0.016;
Matches 68; Conservative 38; Mismatches 124; Indels 69; Gaps 14;
QY 32 SLSSAKGRCAVLEPLPDHR---RNLGAVGGGKSNNDYDEPEFOLLKAWPSMKILPA 88
DB 336 SSSLVNGRLHEL--PVPKPRGTGTPKSEGTPAPRK-----DPPWITLVQAEPKKK--- 382

```
Qy 89 RPIQSEYADTRYFQDMMEAPLLPPKASVSTERQ--TRDVRMTOLEEVDKPT-FKDVRSQ 146
Db 383 -----PAPLPSSSPSPSPSQDSROVENGTEVAQSPSTASLESK 422
Qy 147 RFKGFYTKINKTLPPLPPAI--TLPKYQPLPAPPESSAYFAPKPT-PPEVORGRQ 204
Db 423 PYNFEEEDKEEBAAPASLATSPALGH--PESTPKSLHPWYGITPTSPKTKKRPAP 480
Qy 205 RSKADFRLVGAEEESHOTKPESSCFPS-----SNQNTOKS-----PPAIASSY 249
Db 481 RAPASPLALHASRLSH--SEPPSATPSALSVESLSSESASQAGAELEPPAVPKSSS 538
Qy 250 MPQKH-----SIQARDHTGSMQHCPAQRCAAAASHSPMLPYENTNSEKPDPTKP 299
Db 539 EPAVHAPGTGPNVSLSTNSGLASSGELVPRVQMPQASPLGLAP-RTRGSSGQPAPK 596

RESULT 9
US-10-330-773-627
; Sequence 627, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 627
; LENGTH: 2344
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-330-773-627

Query Match 5.2%; Score 121.5; DB 6; Length 2344;
Best Local Similarity 19.9%; Pred. No. 0.18;
Matches 81; Conservative 45; Mismatches 151; Indels 131; Gaps 17;

Qy 46 PLPDHRRNLGVPGEKCNNDYEDPEFQLLKAWPSMKILPAPPIQSEYADTRYFQDM 105
Db 157 PAKSMWNKQGGQGGIQVNS-----OPQOEFPSSLQAAGDOEKKEKANDENY---- 203
Qy 106 MEAPLLLPK-----ASVSTERQTRDVRMTOLEEVDKPTFKDVRSQRFKGF 151
Db 204 GPGLSLPPNVACHRDGKSNAGSSSDODEKQLQODESTAITSEONDILKVKVEKRIACGP 263
Qy 152 KYTKIN-----KTELPP-----PRPAITLPKKYQPL--PPAPPESSAY--FAPKPTF 195
Db 264 POAKLNGQQPASQYRAMMPPYMFQYPRMAYPPLHGPMPFPPLSLSEANKSLRGRGPPPSW 323
Qy 196 PEVORGRQRSK-----DPSRVLG--AEESHHTKTPES 228
Db 324 ASEPERSEILSASELKBELDKFDNLDADEAGWAGAEVNDYTEQLNFSDDDEQGSTSPKE 383
Qy 229 SCPSNNQNTOKSP-----PAIASSYMPGKHSIQARDHTGS 264
Db 384 S--SSEDTAKTPESTENRKEVDEASTKSSQIPAPQPPVTKSPYKGGPPFNQER---GP 438
Qy 265 MQHCPAQRCAAAASHSP---RMLPYENTNSEKPDPTK---PDEKDVW----- 305
Db 439 SSSLPPPPKLLAQHPPPPDRQIP-----GRQGPFPSPKPPVPDNDENIWKQRRKQOSEISAA 494
Qy 306 -----QNEWYGEYSRQAVEVDVLMKENKDGTFILVRCSTKSKAEPYLVVFG- 358
Db 495 VERARKREBERMERQKAAQDSNRSE-KETTQVVQEAEPESGAQP 541

RESULT 10
US-11-221-263-24
; Sequence 24, Application US/11221263
```

```
; Publication No. US20060051823A1
; GENERAL INFORMATION:
; APPLICANT: The United States of America as Represented by the Secretary of the
; APPLICANT: Department of Health and Human Services, Centers for Disease Control and
; APPLICANT: Prevention
; APPLICANT: Liu, Hsi
; APPLICANT: Steiner, Bret
; APPLICANT: Berta, Rodes
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING TREPONEMA PALLIDUM
; FILE REFERENCE: 6395-61666
; CURRENT APPLICATION NUMBER: US/11/221,263
; CURRENT FILING DATE: 2005-09-06
; PRIOR APPLICATION NUMBER: US/10/017,168
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: PCT/US00/16425
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/138,981
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Treponema pallidum
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Subspecies: endemicum (Bosnia strain)
US-11-221-263-24

Query Match 5.1%; Score 117; DB 7; Length 428;
Best Local Similarity 20.4%; Pred. No. 0.047;
Matches 79; Conservative 45; Mismatches 163; Indels 100; Gaps 15;

Qy 52 RNLGVPGEKCNNDYEDPEFQLLKAWPSMKILPAPPIQES-----EYADT 99
Db 63 RTLGTVRGSGTSDG-----LSLASLPSR--VPARPAQRDPLSSPPAGHTVPEYRDT 113
Qy 100 RYFQD-----MMEAPLLLPKAS--VSTERQTRDVRMT-----QLEEVDPK 138
Db 114 VIFDDPRLVSPLSREVEDVPKVVPEASEREGGEREVEDVPKVVPEASEREGGEREVEDVP 173
Qy 139 TFKDVRSQRFKGYTKINKTLPPLPPPAITLTKKYQPLPAPPESSAYFAPKPTPEV 198
Db 174 KVVEPASERG-----EREVEDVPKVVPEASEREGGEREVEDVPKVVPEAS 220
Qy 199 QRGPQRORSKDFSRVL-----GAEESHHTK-----PESSCFSPSNQNTOKSPPAIASS 248
Db 221 EREGGEREVEDVPKVVPEASEREGGEREVEDVPKVVPEASEREGGEREVEDVPKVVPEAS 280
Qy 249 YMPGKHSIQARDHTGSMQHCPAQRCAAAASHSPMLPYENTNSEKPDPTKPD--EKDVWQ 306
Db 281 EREGGEREVASQHTKOPSH-----SVSNSAPNQ--FRNPEGELPF-TLPDLSESEIVV 330
Qy 307 NEWYIGEYSRQAVEVDVLMKENKDGTFILVRCSTKSKAEPYLVVFG-----NKVYN 358
Db 331 PESQKRAHPQVPEGAPRQGPGEYVQVAVFHDAIQVOSIVHRYGVETPIAVEQDIHE 390
Qy 359 VKIRF-----LESNQOF 370
Db 391 GKVRFTVCVPQKDERGAVLENFQRF 417

RESULT 11
US-10-821-231C-1
; Sequence 1, Application US/10821231C
; Publication No. US20050275837A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Dongmano
; APPLICANT: Ben-Amotz, Dor
; APPLICANT: Xie, Yong
; APPLICANT: Davison, Vincent J.
; APPLICANT: Mrozek, Melissa
; APPLICANT: Ortiz, Corabi
```


This Page Blank (uspto)